Development of an Integrated Expression Platform for Protein Production in Eukaryotic Hosts

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Steffen Meyer

aus Bielefeld

1. Referent:	Professor Dr. Stefan Dübel
2. Referent:	Professorin Dr. Christiane Ritter
Eingereicht am:	7. Mai 2012
mündliche Prüfung (Disputation) am:	16. Juli 2012

Druckjahr 2012

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge und Seminarvorträge

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. A New Versatile Donor Vector for Protein Production in Eukaryotic Systems, SGC Workshop on Human Protein Production, University of Oxford (2011)

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. A New Versatile Donor Vector for Protein Production in Eukaryotic Systems, PEGS Europe - Protein & Antibody Engineering Summit, Hannover (2011)

Meyer, S., Lorenz C., Sprick, G., Baser, B. & van den Heuvel, J. A New Versatile Donor-/Expression Vector for Protein Production in Eukaryotic Systems, Animal Cell Technology Industrial Platform (ACTIP) Biannual Meeting, Lyon (2011)

Meyer, S., Lorenz C., Sprick, G., Baser, B. & van den Heuvel, J. mHost-XS - A versatile System for Protein production in eukaryotic systems, Seminar on special topics of Molecular and Technical Biochemistry, Technische Universität Braunschweig (2011)

Posterbeiträge

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. New Versatile Donor Vectors for Protein Production in Eukaryotic Systems, ISBioTech 1st Annual Meeting, Norfolk, VA (2011)

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. A New Versatile Donor Vector for Protein Production in Eukaryotic Systems, PEGS Europe - Protein & Antibody Engineering Summit, Hannover (2011)

Content

Abbreviations

aa	Amino acids
AcMNPV	*Autographa californica* multicapsid nucleopolyhedrovirus
(m)AB	(monoclonal) Antibody
Amp	Ampicillin
ATG	Translation start codon
bp	Base pair
BCIP	5-Bromo-4-chloro-3-indoxyl phosphate
BEVS	Baculovirus expression vector system
Bluo-Gal	5-Bromo-3-indolyl-β-D-galacto pyranoside
BSA	Bovine serum albumine
BV	Budded virus
CHO	Chinese Hamster Ovary
Cm	Chloramphenicol
CMV	Cytomegalovirus
ddH2O	H_2O bidest
DNA	Deoxyribonucleid acid
dNTP	Deoxy ribonucleosid-triphosphate4
ε	Extinctions coefficient [$l \times g^{-1} \times cm^{-1}$]
EBNA1	Epstein-Barr Nuclear antigen 1
EBV	Epstein-Barr virus
E. coli	*Escherichia coli*
ECD	Extra cellular domain
EDTA	Ethylen diamine tetra acetic acid
EF1-α	Elongation factor 1 α
eGFP	Enhanced green fluorescent protein
EtOH	Ethanol

Fc	fragment crystallizable region of an antibody
FCS	Fetal calf serum
Flp	Flippase
for	Forward (-primer)
FRT	Flp recombinase target
FSC	Forward Scatter
Fv	variable fragment of an antibody
Gm	Gentamicin
HEK	Human Embryonic Kidney
His_n	n consecutive Histidin residues as affinity tag
hpi	Hours post infection
hpt	Hours post transfection
IMAC	Immobilised metal ion affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	Kanamycin
kDa	Kilodalton
L	Liter/ lysate
LB	Lysogeny broth
LMP-Agarose	Low melting point agarose
M	Marker (size standard) / Molarity [mol/L]
MALDI	Matrix assisted laser desorption/ionisation
MCS	Multiple cloning site
MeOH	Methanol
MOI	Multiplicity of infection
MS	Mass spectrometry
MW	Molecular weight
MWCO	Molecular weight cut off
NBT	Nitro blue tetrazolium chloride
Neo	Neomycin phosphotransferase

Ni-NTA	Complex of nickel ions and nitriloacetic acid
NPV	Nucleopolyhedrovirus
o/n	Over night
OB	Occlusion body
ODV	Occlusion derived virus
ORF	Open reading frame
oriP	Origin of replication
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PGK	Phosphoglycerate kinase
PI	Propidium iodide
pI	Isoelectric point
polyA	Polyadenylation signal
PVDF	Polyvinylidene difluorid
rev	Reverse (-primer)
RMCE	Recombinase-mediated cassette exchange
rpm	Revolutions per minute
RT	Room temperature / Recombinase target
sc	single chain
SDS	Sodium dodecyl sulphate
Sf	*Spodoptera frugiperda*
SOC	Super optimal broth for catabolite repression
SSC	Sideward scatter
SV40	Simian Virus 40
TAE	Tris-acetate-EDTA
TBS-T	Tris buffered saline with Tween-20
TE	Tris-EDTA

TEMED	N,N,N´,N´-Tetramethylethylendiamin
Tet	Tetracyclin
Tm	Melting temperature
Tn7	Bacterial transposon Tn7
TOF	Time of flight
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
U	Catalytical units
UV	Ultraviolet
v/v	Volume fraction (volume per volume)
w/v	Mass concentration (weight per volume)
×g	Gravitational acceleration
YFP	Yellow fluorescent protein

Synopsis

Eine Grundvoraussetzung für strukturbiologische und biochemische Analysen ist die Verfügbarkeit von Proteinen in großen Mengen und ausreichender Qualität. Die Identifizierung von exprimierbaren Varianten eines Zielproteins und die Evaluierung des am besten geeigneten Expressionswirtes ist jedoch ein zeit- und arbeitsaufwendiger sowie kostenintensiver Prozess. Daher stellt die rekombinante Produktion von Zielproteinen in eukaryotischen Expressionssystemen einen entscheidenden Flaschenhals in der strukturbiologischen Forschung dar.

In dieser Arbeit wurde der neuartige Expressions- und Donorvektor pFlpBtM konstruiert, der für die Produktion von Zielproteinen in den verschiedenen eukaryotischen Expressionssystemen der Helmholtz Protein Sample Production Facility (PSPF) eingesetzt werden kann. Dies wurde durch die bislang nicht verfügbare Kombination der Elemente mehrerer leistungsfähiger Expressionsmethoden in einem einzigen Plasmid ermöglicht. Der Vektor eignet sich dadurch sowohl für die parallele Evaluierung von verschiedenen Wirtssystemen sowie für ein schnelles Screening von verschiedenen Proteinkonstrukten mittels transienter Expression. Die Funktionalität des Vektors konnte über die erfolgreiche Expression unterschiedlicher Modellproteinklassen in der humanen Zelllinie HEK293-6E, dem Baculovirus Expressionsvektorsystem sowie in stabiler genomischer Expression nach Rekombinase-vermitteltem Kassettenaustausch (RMCE) in einer dedizierten CHO Lec3.2.8.1 Masterzelllinie gezeigt werden.

Die Expressionsplattform wurde zudem um speziell entwickelte Insektenzelllinien erweitert, die eine stabile genomische Expression von integrierten Proteinen ermöglichen und sich daher als optimierte Wirtszelllinie für die baculovirale Expression von besonders anspruchsvollen Zielproteinen eignen. Die bislang nicht verfügbare Implementierung des flexiblen RMCE für die Integration von auxiliaren Proteinen in Insektenzellen ist kompatibel zu dem bereits vorhandenen System in CHO Lec3.2.8.1 Zellen. Durch die Kombination des vielseitigen Vektors und optimierten Wirtszellen bietet die PSPF jetzt eine integrierte Infrastruktur zur Proteinproduktion in eukaryotischen Systemen, die als „multi-host expression system“ (mHost-XS) bezeichnet wird.

Synopsis

A prerequisite for structural and biochemical analyses of proteins is the availability of large amounts of high-quality material. However, the recombinant production of complex proteins typically requires a profound screening process to identify expressible constructs. Furthermore, the evaluation of an optimal expression host has an important impact on protein quality, yield and cost efficiency. As a result, the production of target proteins in eukaryotic systems remains a major bottleneck in structural biology.

In this work, the novel expression/donor vector pFlpBtM has been constructed to be utilised for both, the rapid screening for expressible protein variants and the evaluation of an optimal expression host. The range of applications of this vector comprises a variety of highly efficient eukaryotic hosts for recombinant protein production used at the Helmholtz Protein Sample Production Facility (PSPF). It was realised by the unique combination of the elements required for different expression methods including transient, viral and stable expression strategies. The efficiency of the vector is demonstrated in a comparative expression study of model proteins from different protein classes. In order to optimise the production strategy for each protein, they have been produced employing pFlpBtM as the donor- or expression vector by transient expression in HEK293-6E cells as well as in the baculovirus expression vector system and by stable expression using a recombinase mediated cassette exchange system in CHO Lec3.2.8.1 cells.

Additionally, engineered BEVS host cell lines were generated, which enable the stable genomic expression of auxiliary proteins to facilitate the production of particularly challenging target proteins. The flexible RMCE system was implemented in insect cell lines for the first time to provide optimised coexpression for each target protein and deploy compatibility to the RMCE CHO Lec3.2.8.1 cell lines previously developed in our group. With the combination of a versatile donor/expression plasmid and optimised mammalian and lepidopteran producer cell lines, the PSPF offers an integrated platform for protein production in eukaryotic systems. This high performance infrastructure was therefore termed “multi-host expression system” (mHost-XS).

1 Introduction

1.1 Recombinant protein expression

Proteins represent the most versatile class of biological molecules since they are involved in virtually every essential process within cells, such as metabolism, regulation, signalling, transport or structuring. Thus, knowledge of structure and function of proteins is indispensable for the understanding of biological processes and concepts. Moreover, information about functional sites and binding partners of proteins that are involved in pathogenesis substantially facilitates the development of drugs and treatment strategies (Buchanan, 2002).

The beginnings of protein research date back to the early 19th century. In the paper "On the composition of some animal substances" Gerardus J. Mulder described molecules much larger than those chemists of the time were used to deal with (Mulder, 1838). His hypothesis that there was a fundamental substance from which all proteins were composed, led to the identification of the amino acids (Perrett, 2007). About 200 years later, proteins are not only the aim of basic research, but also have become a versatile tool for biotechnological and pharmaceutical applications. They are used as enzymes in biocatalytical processes, vaccines or therapeutics. This fact illustrates how visionary Jöns J. Berzelius had been, when he coined the term protein in 1838 from the Greek word πρώτειος, meaning "primary" to emphasise its importance.

However, a prerequisite for the structural and functional analysis of proteins as well as for their use in biotechnological or pharmaceutical applications is the availability of proteins in large quantities and high quality. These demands usually cannot be met by isolating proteins from natural sources. Thus, protein production is performed by recombinant expression, which not only allows to comply with the requirements mentioned before, but also enables the generation of genetically optimised products or isolated protein domains that are not available *in vivo*. These include fusions with markers for identification, purification or improved solubility (Arnau *et al.*, 2006), as well as truncated protein variants, single domains, hybrid proteins or novel engineered protein species such as alternative antibody formats (Carter, 2006; Kontermann, 2011).

Heterologous expression of proteins was not possible until the development of recombinant DNA technology in the early 1970s (Cohen *et al.*, 1973). Rapid progress in that field enabled the recombinant production of human insulin by the end of that decade (Goeddel *et al.*, 1979) that finally was available as the first biotechnologically produced pharmaceutical by Eli Lilly in 1982. Ongoing advances over the last 30 years allowed an extension of the spectrum of expression hosts beyond prokaryotic systems, thus today also yeasts as well as invertebrate or vertebrate cell lines are used for the heterologous production of proteins. The diverse hosts show major differences regarding the required expression strategy as well as in protein quality and yield (Hunt, 2005). Furthermore, protein production in eukaryotic cell lines is generally more time-consuming and expensive than in bacteria. Thus, a profound screening for the best protein construct as well as the most appropriate host regarding both yield and quality of protein is essential. Accordingly, the development of new methods and optimised production procedures represent a worthwhile challenge for biotechnological research.

1.2 Bacterial and fungal expression hosts

E. coli is the predominant host for recombinant protein expression due to its well-characterised genetics and its ability to grow rapidly at high densities on inexpensive substrates (Baneyx, 1999). Today, a large variety of optimised dedicated expression strains and an increasing number of cloning vectors are available (Sorensen & Mortensen, 2005). Furthermore, cultivation techniques have been established that permit labelling of recombinant proteins with heavy atom derivatives such as selenomethionine or with stable isotopes such as ^{2}H, ^{13}C and ^{15}N facilitating structural studies by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy (Bernard & Payton, 2001). Besides *E. coli*, strains of the gram positive *streptomyces* and *bacillus* species are also used as bacterial expression hosts. *B. megaterium* for instance is distinguished by the absence of endotoxins and its efficient secretory pathway. Thus, recombinant production of secretory proteins such as Penicillin G amylase or antibody fragments has already successfully been performed in this host (Yang *et al.*, 2006; Jordan *et al.*, 2007). However, many viral and mammalian proteins either require post-translational modifications for proper folding and full biological activity or have to be coexpressed with binding partners as a multi protein complex for their structural stability (Nettleship *et al.*, 2010). Since

prokaryotes are not capable to impart these modifications, misfolding of heterologous proteins may occur upon overexpression leading to the formation of inclusion bodies (Gray, 2001). As a result, only 10 - 20 % of eukaryotic proteins can be expressed in *E. coli* in soluble form (Büssow *et al.*, 2005). Therefore, eukaryotic expression hosts are indispensable for recombinant production of those proteins and consequently about 10 % of all protein structures submitted to the Protein Data Base (PDB) have been solved after expression in eukaryotic systems.

Besides mammalian cell cultures and baculoviral expression in insect cells which are described more detailed in chapters 1.3 and 1.4, yeasts are the third major eukaryotic system for recombinant protein expression. As single-cell eukaryotic hosts, they combine some of the advantages of prokaryotic and eukaryotic based expression systems being physically robust and amenable to high-density fermentation and additionally possess the necessary cellular machinery to carry out post-translational modifications (Aricescu *et al.*, 2006). The methylotrophic yeast *Pichia pastoris* and the baker's yeast *Saccharomyces cerevisiae* are the most important fungal expression hosts. In *P. pastoris*, more than 400 different heterologous genes have been expressed under the control of the strong and tightly regulated *P. pastoris* alcohol oxidase 1 (AOX1) promoter (Cereghino *et al.*, 2002). The combination of strong expression and efficient secretion mediated by the MATα prepro secretion signal (Brake, 1990) allows to produce large amounts of recombinant proteins, which accumulate in the culture supernatant (Cregg *et al.*, 2000). Additionally, minimal media enable the production of ^{13}C-labelled proteins for NMR-based structural analyses (Laroche *et al.*, 1994). The major disadvantage in the use of *Pichia* as expression host is that no episomal replication of the expression plasmids is possible. Transfer vectors bearing the gene of interest need to be integrated into the host cell genome via homologous recombination. The time-consuming screening for a expression clone limits the use of this system for a parallel high-throughput screening. In *S. cerevisiae* genes of interest are expressed under the control of the copper inducible metallothionein promoter (CUP1). Although episomal replication of genetic elements harbouring the transgene is possible in this host, the quality of recombinantly produced target proteins is adversely affected by the occurrence of hyperglycosylation. Proteins expressed in *S. cerevisiae* often carry poly-mannose glycans composed of up to 350 mannose residues (Nakamura *et al.*, 1993) and are thus not suitable for crystallisation as complex and inhomogeneous glycan side chains inhibit the formation of protein crystals (Baker et al., 1994).

1.3 Baculoviral expression in insect cells

Lepidoptera cell lines are favourable for the production of proteins for crystallization due to their simple glycosylation pattern (Tomiya *et al.*, 2004) and the scalability of the virus-dependent expression in suspension culture. Hence, the Baculovirus Expression Vector System (BEVS) is predominantly used for the production of proteins for structural analyses with a share of almost 50 % among all eukaryotic systems (Figure 1-1). Since Lepidoptera are the natural hosts for *baculoviridae*, infection, viral replication and promoter activity are highly effective. Thus, yields of up to 75 mg recombinant proteins per L have been reported (Meghrous *et al.*, 2009).

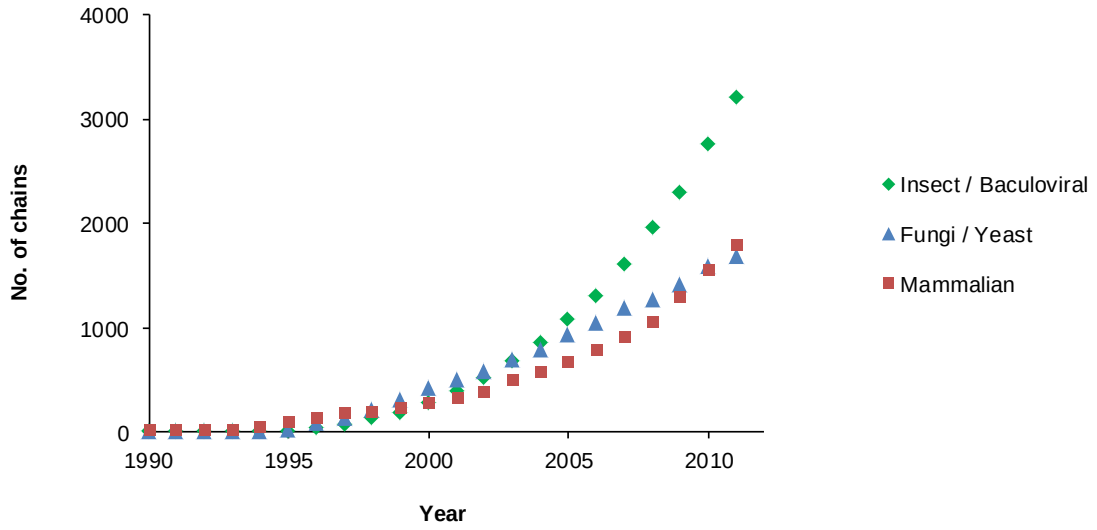


Figure 1-1: Total number of protein chains deposited in the PDB by expression host
Cumulative total number of protein chains in the PDB whose expression system was identified as mammalian, insect, baculovirus, fungi or yeast is plotted by year of deposition. Expression data were parsed by Joachim Reichelt from the set of PDB files available as of December 2011. Chains were counted rather than PDB entries as expression information is recorded by chains in the PDB.

Three lepidopteran cell lines are of major relevance for recombinant protein production in the BEVS. IPLB-SF-21AE (Vaughn *et al.*, 1977) was generated at the USDA Insect Pathology Laboratory (IPLB) from pupal ovarian tissue cells of the Fall Armyworm *Spodoptera frugiperda* (Gardiner & Stockdale, 1975). In 1983 the Sf9 cell line was derived as a subclone thereof by G. Smith at Texas A&M University (Summers & Smith, 1987). In recent years, genetically engineered Sf9 subclones have been described, where stable genomic integrations of mammalian glycosyltransferases enable their use for the

production of glycoproteins with human like glycosylation (Jarvis *et al.*, 1998; Hollister *et al.*, 2002; Legardinier *et al.*, 2005). The third important host cell line is BTI-Tn-5B1-4 (Wickham & Nemerow, 1993) which had been generated at the Boyce Thomson Institute in Ithaca and is commercialised by Invitrogen as High Five™. Its parental line BTI-Tn-5B1-28 had been isolated from ovaries of the Cabbage Looper *Trichoplusia ni* (Granados *et al.*, 1986). Other less common host cell lines for protein production are MB0507 derived from *Mamestra brassicae* (King *et al.*, 1991) or Bm5 from *Bombyx mori* (Grace, 1967).

While the simple and homogenous glycosylation pattern provided by insect cell lines is advantageous for producing proteins for structural biology, it is likewise a major drawback for the production of pharmaceuticals. The high-mannose and paucimannose type glycan chains would compromise *in vivo* bioactivity of therapeutic proteins and potentially induce allergenic reactions (Durocher & Butler, 2009). However, as a viral system it is well-suited for producing antigens and virus like particles for the development of vaccines. The first licensed vaccine produced in insect cells is Cervarix™ by GlaxoSmithKline. It is a truncated form of the human papilloma virus major capsid protein L1 types 16 and 18 and is administered to prevent cervical cancer. By now, more than 10 BEVS-derived products are now either commercially available or in clinical trials, e.g. vaccines against prostate cancer (Provenge™) as well as seasonal and pandemic influenza (Flublok™, Chimigen™) and some veterinary vaccines (Cox & Hollister, 2009; Durocher & Butler, 2009; Krammer & Grabherr, 2010). Additionally, the production of insect cell derived virus like particles of *filoviridae* like Marburgvirus and Ebolavirus is in early stage development (Ye *et al.*, 2006).

1.3.1 *Autographa californica* multicapsid nucleopolyhedrovirus

First descriptions of baculoviruses date back almost 200 years, when silk worm diseases, which represent a large economic risk to the silk industry, were investigated by the upcoming light microscopy techniques. The presence of highly refractile occlusion bodies that were symptomatic of the affected insects could be characterised as a prominent feature. These polyhedron shaped particles led by the mid-1800s to the naming of the diseases associated with these structures as 'polyhedroses' (Figure 1-2). However, it took another hundred years until the availability of high-resolution electron microscopy in the late 1940s before the presence of rod-shaped baculoviral virions within the occlusion bodies (OB) was demonstrated (Bergold, 1947).

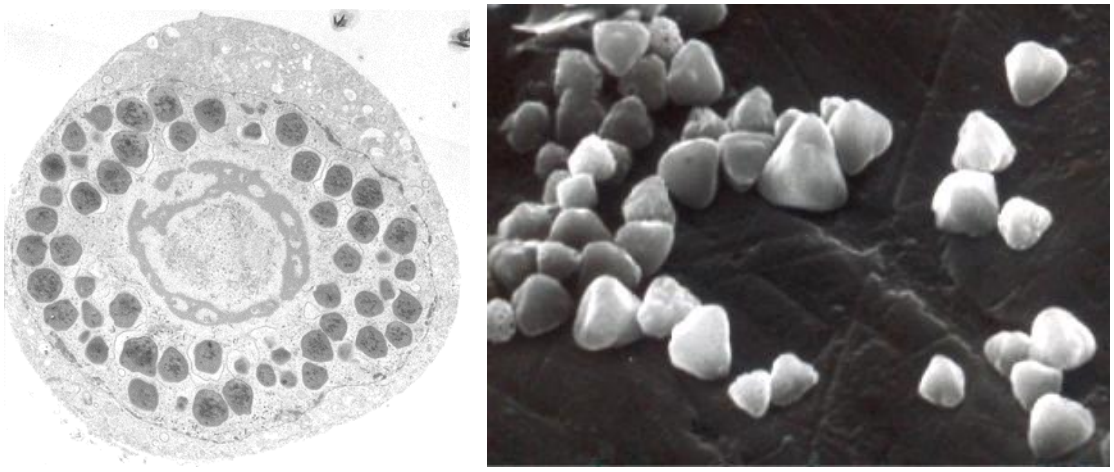


Figure 1-2: Baculoviral occlusion bodies
Occlusion bodies are crystalline protein matrices containing numerous virions. They are formed in the nucleus upon hyperexpression of the viral proteins polyhedrin and P10 during the late phase of the infection. The left picture shows a cross section of an infected cell prior to cell lysis (picture courtesy of Department of Chemical Engineering, The University of Queensland). The right picture is an EM scanning of released polyhedral occlusion bodies upon disintegration of the host (source: A: K. Hughes and R. B. Addison taken from Rohrmann, 2008).

The family of the *baculoviridae* belongs to the double-stranded DNA viruses and is divided into the genera of *Granulovirus* and *Nucleopolyhedrovirus*. The latter can be further divided into single-nucleopolyhedroviruses (SNPV) and multiple-nucleoplyhedroviruses (MNPV) depending on the number of nucleocapsids per virion. The *Autographa californica* multicapsid nucleopolyhedrovirus (*Ac*MNPV), which is used as a vector in the BEVS is infective for more than 30 different Lepidoptera species. As a member of the MNPVs, a virion consists of several nucleocapsids, each containing a copy of the 134 kbp large circular genome (Ayres *et al.*, 1994). The polyhedral occlusion bodies are formed by embedding numerous of these virions in a matrix of the polyhedrin protein (Figure 1-3 A). The immobilization of the infective particles within the crystalline protein lattice provides an environment where they are protected from extreme temperatures and UV light. This allows the virus to remain viable indefinitely outside of the host (Rohrmann, 2008).

Upon ingestion by the host, the alkaline pH in the midgut of the insect causes the dissolution of the polyhedrin matrix of the OB. The released occlusion derived viruses (ODV) subsequently cause a primary infection of the epithelial cells (Figure 1-3, B). By fusion of the virions with the peritrophic membrane the nucleocapsids enter the cells (Volkman & Goldsmith, 1985). In the cytosol their trafficking is mediated by actin polymerization by which they finally reach the nucleus (Ohkawa *et al.*, 2010).

Approximately 6 h post infection (hpi), the first replicated viruses are released by budding from the membrane as budded virus (BV). They spread the infection via the hemolymph to the deeper tissue; a process called secondary infection, which finally leads to a systemic infection of the entire host.

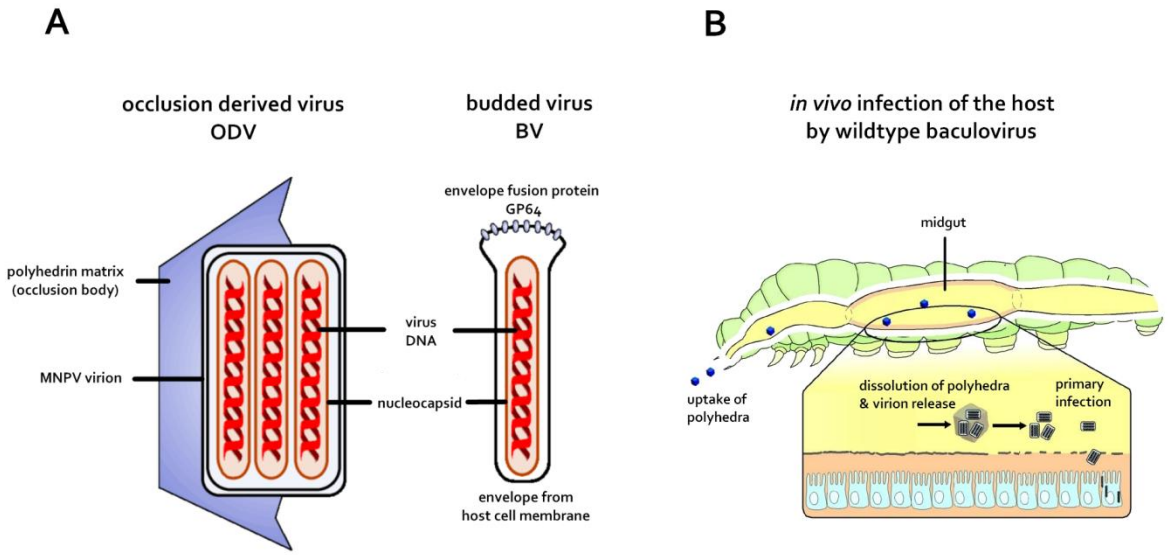


Figure 1-3: Schematic view of the *Ac*MNPV and the infection of the host *in vivo*
A: In MNPVs a virion contains multiple nucleocapsids. Outside of the host numerous virions are occluded by a proteinacious matrix formed by the viral proteins polyhedrin and P10. Secondary infection of deeper tissue in the host as well as in cell culture is mediated by budded viruses. The nucleocapsid is enveloped by a membrane during budding from the host cell. The viral protein gp64 is presented at the surface and enhances the fusion with cell membranes thereby mediating virus entry. **B:** Primary infection occurs after the uptake of occlusion bodies by the insect larvae and dissolution of the protein matrix in the midgut (edited work, original source: A. J. Cann, www.microbiologybytes.com).

However, a significant part of the newly assembled nucleocapsids are not released as BV, and remain in the nucleus instead. These particles are occluded in the late phase of the infection. This process starts upon expression of very late genes resulting in the production of high levels of the matrix proteins polyhedrin and P10 24 -96 hpi. Finally, viral proteases and chitinases cause the liquefication of the host and degradation of the chitinous exoskeleton, thereby releasing the OB into the environment (Figure 1-4).

Figure 1-4: Liquefied insect larvae as a result of baculoviral infection
In the very late phase of the systemic infection, the expression of the protease cathepsin and the chitinase Ac126 lead to a total disintegration of the host. This process is termed melting or liquefaction and permits the release of occluded virions to the environment (source: L. King, Oxford).

1.3.2 Baculovirus expression vector system

The first biotechnological application of *baculoviridae* aimed at agricultural purposes. Intensive studies in the 1970s enabled the success of using a nucleopolyhedrovirus to control the Douglas-fir tussock moth (Martignoni, 1984). Subsequent effort was taken to understand the molecular biology of *baculoviridae*. It was shown that BV are up to 1,000-fold more efficient at infecting cultured cells compared to ODV, which have a 10,000-fold higher ability to enter midgut epithelial cells instead (Volkman *et al.*, 1976; Volkman & Summers, 1977). Moreover, the matrix proteins polyhedrin and P10, which are hyperexpressed in the late phase of infection, are not essential for viral reproduction and infectivity in cell culture (Kuzio *et al.*, 1984; Williams *et al.*, 1989; Summers, 2006). These findings directly promoted the use of *Ac*MNPV for recombinant protein expression *in vitro*. The BEVS was pioneered at the laboratory of Max D. Summers at the Texas A&M University and patented in 1988 (US Patent No. 4,745,051).

The first version of the BEVS made use of homologous recombination of transgenes into the polyhedrin locus. Therefore host cells were cotransfected with viral DNA and a donor plasmid on which the gene of interest is flanked by corresponding sequences. Since the double crossing over essential for the integration is a very rare event (Glick *et al.*, 2009), only 0.1 % of the viruses carry the gene of interest. Therefore, a purification of recombinant virus based on a phenotypic evaluation of the cells was necessary: those cells

in which the polyhedrin gene is lacking due to the exchange by the gene of interest do not exhibit the formation of occlusion bodies. A first improvement was established by the integration of a unique *Bsu*36I restriction site into viral genome. This enabled the transfection of linearised viral DNA which is not replicated in the cells unless recircularised upon homologous recombination with the donor vector. This improvement increased the yield of recombinant viruses from 0.1 % to 30 % (Kitts *et al.*, 1990) and is still in use in commercial available systems like BaculoGold™ (BD) or BacPAK™ (Clontech) with efficiencies up to 99 %. A similar system relies on the deletion of essential gene sequences within open reading frames (ORFs) of the viral genome. Only by restoring the gene functions by homologous recombination, recombinant virus retrieves its ability to replicate (Kitts & Possee, 1993). It is commercialised by Novagen as BacMagic™ and achieves also nearly 100 % effectiveness.

The most versatile strategy for the generation of recombinant viruses is shown in Figure 1-5 A and relies on the integration of genes into a baculovirus shuttle vector in *E. coli* (Luckow *et al.*, 1993). It was developed at Monsanto by the integration of DNA elements into the polyhedrin locus that comprise a miniF-replicon, which enables autonomous replication in *E. coli* as a bacterial artificial chromosome (Hosoda *et al.*, 1990; Shizuya *et al.*, 1992). Furthermore, the integration contains a kanamycin resistance gene for selection and the *lacZ* sequence encoding β-galactosidase in which *att*Tn7 sites for the bacterial transposon Tn7 are located (Grinter, 1983). The integration of genes into the bacmid follows a two-step process. Firstly, the gene of interest is cloned into the expression cassette of a donor plasmid comprising a baculovirus promoter and flanking Tn7 sequences. Upon transformation of the *E. coli* strain containing the bacmid the cassette is transposed into the viral genome by a bacterial transposase coded on a helper plasmid (Luckow *et al.*, 1993). The integration disrupts the *lacZ* ORF, thus allowing to identify positive recombinants by blue-white selection. Following isolation of the recombinant viral DNA, it can be directly used for the transfection of insect cells without the need for subsequent plaque purification steps. This principle was commercialised as Bac-to-Bac™ system by Invitrogen in 1994 (US Patent No. 5,348,886) and constituted the basis for novel systems emerged thereof such as MultiBac (Berger *et al.*, 2004) or ACEMBL (Bieniossek *et al.*, 2009; Trowitzsch *et al.*, 2010), which are described in Figure 1-5 B&C.

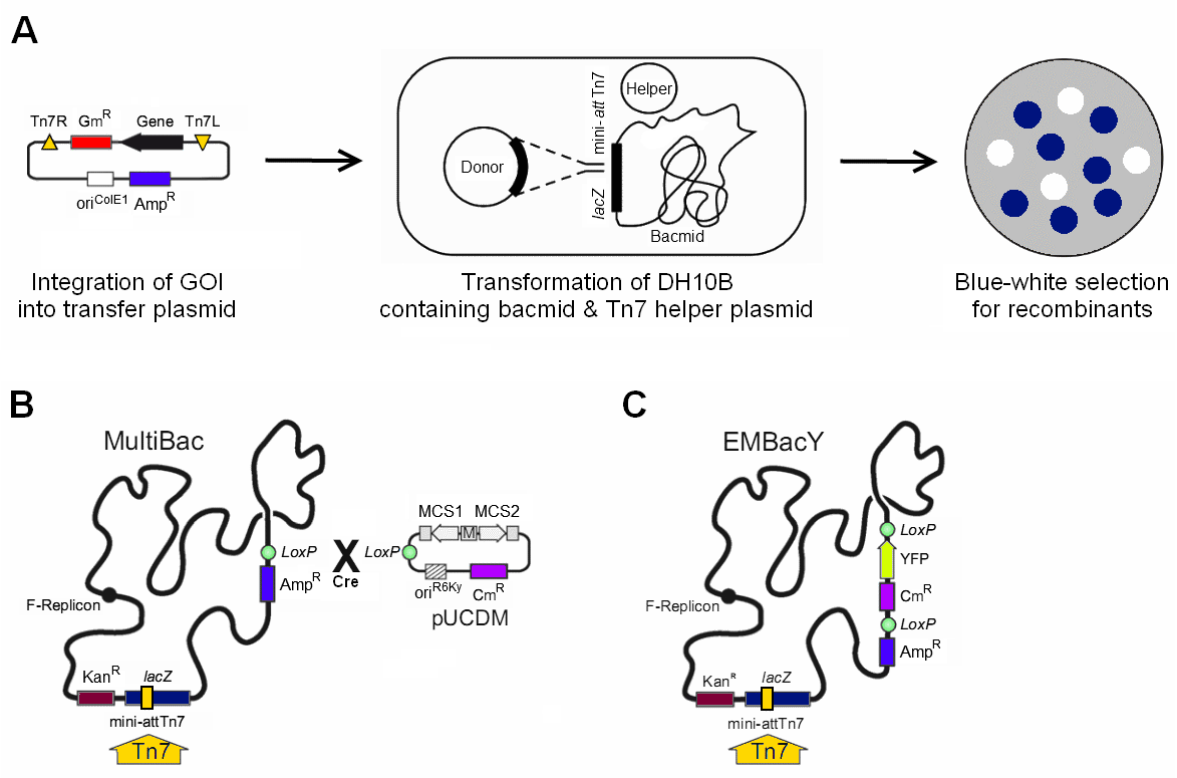


Figure 1-5: Generation of recombinant viral DNA by Tn7 transposition & multigene integration in MultiBac system
A: The gene of interest, present on a transfer vector, is integrated via Tn7 transposition into a baculovirus genome maintained as a bacmid in DH10B *E. coli* cells. Screening for composite bacmids can be performed by blue-white selection, since the Tn7 attachment sites on the bacmid are embedded within a *lacZ* gene. **B:** In the MultiBac system a second integration locus was introduced by a deletion of the baculoviral genes coding for a cathepsin protease (v-cath) and a chitinase (chiA). They have been replaced by a loxP imperfect inverted repeat together with an ampicillin resistance marker. This facilitates the integration of a second donor plasmid via Cre recombination. Due to the special design of both donor plasmids, integration of up to 8 genes is possible for coexpression using a single virus. **C:** A more recent bacmid is EMBacY, which has been constructed by the integration of an expression cassette for enhanced YFP production under control of the polh promoter. It allows for efficient monitoring of virus performance and expression of other heterologous proteins integrated into the Tn7 site from the same bacmid (edited work, original sources: Trowitsch, *et al.* 2010 & Bac-to-Bac manual, Invitrogen).

1.4 Mammalian expression hosts

Protein production in mammalian expression systems provides authentic folding, assembly and post-translational modification of proteins of the same origin. Thus, about 70 % of all proteins for pharmaceutical or clinical applications are produced in mammalian cell culture (Wurm, 2004). In the last 20 years, progress in genetic engineering, high-throughput techniques, media composition and process control has led to a 100-fold increase in volumetric productivity so that yields in large scale cultivation have recently reached the gram per litre range (Andersen & Krummen, 2002; Brondyk *et al.*, 2009) .

Two different strategies are used for the production of recombinant proteins in mammalian cells: transient and stable expression. The selection of an appropriate method is dependent on the characteristics and intended applications of the protein (Brondyk *et al.*, 2009). Whenever criteria such as growth behaviour of host cells, clonality, stability and reproducibility are of major importance, stable expression is favoured. This is the case in the production of proteins for crystallization or biochemical assays as well as for pharmaceutical applications. Thus, most therapeutic antibodies currently on the market are produced in stable CHO cell lines. Conversely, recombinant protein production for research purposes (e. g. in HTP screenings, test expressions or construct evaluation) is mainly driven by the cost-effectiveness, simplicity, and speed of the process. In this case transient gene expression - predominantly performed in HEK293-derived cell lines - is superior (Geisse & Fux, 2009). Although recombinant protein expression using the BEVS is also transient, the term is only used to refer plasmid based expression in this work.

1.4.1 Transient expression in HEK293

The cell line HEK293 originates from a primary culture of human embryonic kidney epithelia, which was transformed by a type 5 adenovirus (Graham *et al.*, 1977). Numerous improvements in culturing techniques led to its dominant use in the production of recombinant proteins by transient gene expression. The development of high-effective as well as cost-efficient transfection methods utilizing calcium-phosphate or polyethylenimine (Graham & van der Eb, 1973; Boussif *et al.*, 1995) and adaptation to serum free culture media (Cote *et al.*, 1998) enabled an upscaling of transient expression in HEK293 to bioreactor scale (Pham *et al.*, 2003). Due to these improvements, yields of several hundred mg per L can be achieved (Baldi *et al.*, 2007) and various proteins have been successfully produced for X-ray structure determinations (Bishop *et al.*, 2009; Forneris *et al.*, 2010; McLellan *et al.*, 2011). To further improve recombinant protein expression in this cell line, genetically engineered variants have been developed. These strains allow for episomal replication of expression plasmids by a genomic integration of the simian virus 40 (SV40) large T-antigen (Lebkowski *et al.*, 1985) in HEK293T and the Epstein-Barr virus (EBV) nuclear antigen 1 (EBNA1) in HEK293-6E, respectively (Van Craenenbroeck *et al.*, 2000; Durocher *et al.*, 2002). EBNA1 is a multifunctional protein consisting of a DNA binding domain, a transcription activating domain and a nuclear localisation signal (Langle-Rouault *et al.*, 1998; Dean *et al.*, 1999; Dean *et al.*, 2005). The

protein binds specifically to the Epstein-Barr-Virus origin of replication (EBV-oriP) and mediates the nuclear transport of cytoplasmic DNA (Ambinder *et al.*, 1991; Fischer *et al.*, 1997). There plasmids harbouring the EBV-oriP are replicated, allowing the dissemination of plasmid copies to the daughter cells upon cell division (Lindner & Sugden, 2007; Kishida *et al.*, 2008). The combination of higher copy numbers, increased nuclear localization, transcription activation and extended plasmid half-life in the cell culture, lead to impressive yields of more than 1 g/L with this system (Backliwal *et al.*, 2008a).

1.4.2 Stable expression in CHO

Stable expression of recombinant proteins requires the integration of the gene of interest into the host cell genome. Thereby it facilitates the reproducible production over extended cultivation periods as well as efficient upscaling in bioreactors with several thousand litres (Xing *et al.*, 2009). Thus, this method is commonly used when large quantities of recombinant proteins are needed. Under optimised conditions yields of up to 5 g/L have been reported (Durocher & Butler, 2009). Likewise, stable expression over a long period of time in perfusion mode is also especially beneficial to gain significant amounts of hard-to-express proteins, which are produced only with extremely low yields. The antihemophilic factor VIII, with 2,332 amino acids an extremely large and fragile protein, is reliably being manufactured in suspension-cultivated baby hamster kidney cells (BHK) and harvested continuously for up to 6 months with yields of 150 g/year (Böedeker *et al.*, 1994).

Since the Chinese hamster ovary cell line CHO (Tjio & Puck, 1958) integrates exogenous DNA with a high efficiency (Hoeijmakers *et al.*, 1987) it has become the predominant mammalian host cell line for stable expression. The recently accomplished sequencing of the CHO genome (Wurm & Hacker, 2011) will probably pave the way for directed metabolic engineering to develop optimised producer strains for various purposes during the next decade. A series of glycosylation mutant cell lines such as CHO Lec3.2.8.1 (Stanley, 1989) are already available for the production of glycoproteins for structural analyses. The proteins derived from CHO Lec3.2.8.1 possess only N-linked glycans of the high mannose type $Man_5GlcNAc_2$ which can be further truncated efficiently to a single N-acetylglucosamine (GlcNAc) by endoglycosidase H (endo H) making them more homogeneous and even more likely to crystallise (Davis *et al.*, 1993). Thus, structures of various glycoproteins produced in this cell line have successfully been solved (Butters *et al.*, 1999; Niemann *et al.*, 2007).

However, the integration of genes into the host cell genome by conventional methods is a rare and random event. Moreover, expression characteristics can dramatically differ within a population of transfected cells due to gene copy variations and neighbouring chromosomal elements that modulate the promoter activity to a high extent *in cis* (West & Fraser, 2005). This position effect is also referred to as mosaicism or variegation (Bestor, 2000; Kito *et al.*, 2002). An intensive screening and evaluation procedure is necessary to isolate a stable, high-producer cell clone. This is typically facilitated by using a vector comprising the gene of interest and a selective marker for the integration (Gurtu *et al.*, 1996). The most common selection methods rely on gaining resistance to an antibiotic agent or on the complementation of deficiencies in essential metabolic pathways in the host cell. The integration of bacterial neomycin phosphotransferase gene for instance mediates resistance to the antibiotic Geneticin (G418) (Southern & Berg, 1982), whereas the complementation of the genes of dehydrofolate reductase (DHFR) and glutamine synthetase (GS) enable cells with a corresponding defect to grow in glutamine-free or hypoxanthin- and thymidine-free media, respectively (Urlaub & Chasin, 1980; Cockett *et al.*, 1990). Upon transfection, cells are cultivated for several passages under selective pressure, allowing only those cells to propagate, in which the transfer vector is stably integrated and the transgenes are expressed constitutively. By increasing the concentration of antibiotic or inhibitors, an enrichment of cells with a very high expression rate or gene copy number is possible, eventually obtaining clones with up to several hundred gene copies and increased product yields (Wurm *et al.*, 1986; Kim *et al.*, 2001).

However, identical copies at different *loci* within the genome exhibit a tendency to recombine and thereby causing chromosomal aberration (Derouazi *et al.*, 2004). Likewise, tandem-repeat integrations at the same locus might induce so called "repeat-induced gene-silencing" (RIGS), what is considered to be a defence-mechanism against retroviruses (McBurney *et al.*, 2002). Additionally, prokaryotic sequences on the transfection vector (e. g. replication origin, selection markers) flanking the transgene might influence its transcriptional activity due to their divergent nucleotide composition (Scrable & Stambrook, 1997). Furthermore, removing of antibiotic selection pressure during cell propagation frequently results in declining protein yields caused by mosaic gene silencing (Liu *et al.*, 2006). These effects underline the need for a clonal isolation of cell lines and a subsequent evaluation of their stability and productivity over numerous

passages (Gellissen, 2005). Thus, cell line development from gene transfer to larger-scale production can still take several months to a year despite the development of robotic devices for high-throughput screenings of millions of individual cell clones.

1.4.3 Cell line development by site specific recombination

An alternative method for integrating transgenes into the host cell genome is site specific recombination. This process utilises heterologous recombinase enzymes to induce the deletion, insertion or inversion of DNA fragments at specific sequences within the genome. It was originally developed to study gene functions as well as regulation and expression characteristics of certain *loci* (Branda & Dymecki, 2004). The most common systems are the Cre system, derived from the *E. coli* phage P1 (Stark *et al.*, 1992), the *Streptomyces* phage ΦC31 integrase (Groth *et al.*, 2000) and the Flippase Flp from *S. cerevisiae* (Wirth *et al.*, 2007). The corresponding enzymes Cre and Flp are tyrosine recombinases, whereas ΦC31 is a serine recombinase. They cleave the DNA backbone via a nucleophilic attack on a phosphodiester bond, exchange the strands of the two DNA double-helices involved, and religate the DNA backbones. This reaction does not involve the hydrolysis or synthesis of DNA (Kolb, 2002).

However, a prerequisite for the generation of producer cell lines with these systems is the preceding integration of specific recombinase target sites (RT) into the host cell genome; a process called "tagging". These short sequences are the 34 bp loxP site (Cre), the 48 bp FRT site (Flp) and the heterotypic sites *attB* (34 bp) and *attP* (39 bp) for ΦC31. The generation of tagged master cell lines is based on random integration and requires effort in screening and evaluation. Though, this tagging process has to be performed only once for any host cell line. Additionally, alternative screening procedures based on fluorescent activated cell sorting (FACS) have been established for this purpose (Qiao *et al.*, 2009). Subsequent to the generation of a tagged master cell line, targeted integration of the gene of interest into the pre-tagged locus and the clonal isolation and propagation of producer cell lines is feasible within less than 3 months.

The ΦC31 integrase mediates recombination of two different RTs , attP and attB, located in the host cell genome and a donor plasmid harbouring the gene of interest, respectively. By the recombination process, the plasmid is integrated into the genome and the RT sequences are altered resulting in flanking sites attL and attR. Since those are not

substrates for the integrase, the recombination is a unidirectional process (Thyagarajan *et al.*, 2001). For cell line development a commercial version of this method is offered by Invitrogen termed Jump-In™. However, the occurrence of pseudo integration sites in the mammalian genome, which may cause unexpected integrations cumulating in genomic instability has been proposed (Thyagarajan & Calos, 2005).

Unlike the RT sites of the ΦC31 integrase, the tyrosine recombinases Cre and Flp mediate the recombination between identical sites based on a double-reciprocal crossover. Both, FRT and loxP sites contain an 8 bp asymmetric spacer sequence flanked by long inverted repeats to which the corresponding enzymes bind to mediate the reaction (Figure 1-6). The spacer sequence is responsible for DNA-DNA base pairing and determines the orientation of the RT (Andrews *et al.*, 1985).

A Cre recombinase target site (loxP)

B Flp recombinase target site (FRT)

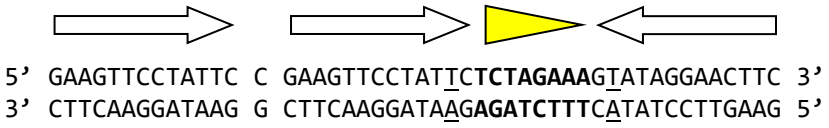


Figure 1-6: Cre target site loxP and Flp recombinase target site FRT
A: The Cre recombinase target site is termed loxP and consists of two 13 bp inverted repeats (light arrows) flanking an 8 bp asymmetrical core, or spacer (blue) in which DNA breaking and rejoining occurs. **B:** The Flp recombinase target site is termed FRT. It also consists of an asymmetric 8 bp spacer (blue) which is flanked by three 13 bp inverted repeats representing three binding sites for Flp (light arrows). Only the two inverted repeats directly flanking the spacer sequence are required for a minimal FRT site. The third upstream repeat, separated by a single base, is present in the 2 µ plasmid but is dispensable for Flp-mediated recombination. Thus, the minimal FRT has a size of 34 bp. By contrast to a loxP site, the inverted repeats of an FRT site contain a single bp difference (underlined).

Cre and Flp can mediate an excision, insertion or a translocation of DNA sequences. The relative orientation of a pair of RT sites to each other determines the type of recombination (Branda & Dymecki, 2004). When two copies of an RT site are arranged as direct repeats, the corresponding enzyme excises the DNA sequence between the copies.

The reverse reaction leads to an insertion of circular DNA containing one RT site. Due to thermodynamic reasons, the entropy-driven, monomolecular excision is favoured to the bimolecular insertion (Baer & Bode, 2001). Flp and Cre mediate an inversion of a DNA sequence to a dynamic equilibrium, when flanking RT sites are positioned in inverse orientation, whereas a recombination of single RTs on two linear DNA molecules results in the translocation of distal sequence (Figure 1-7).

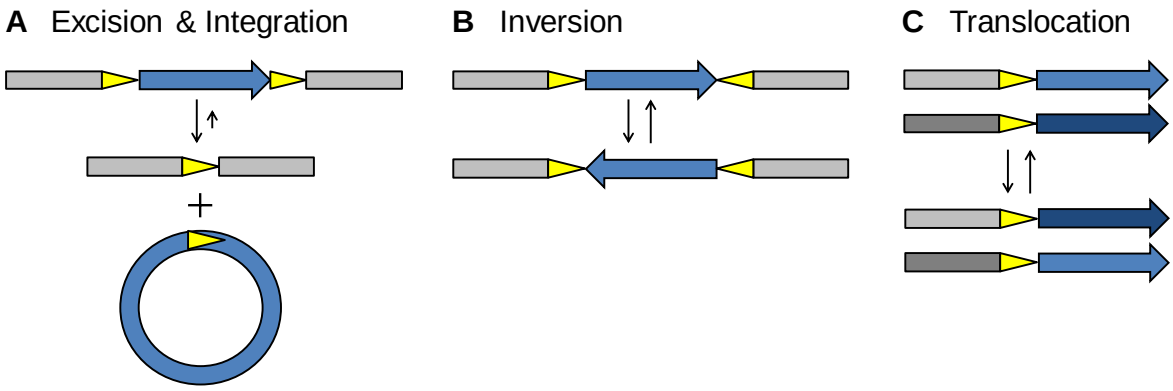


Figure 1-7: Types of recombination mediated by Flp and Cre depending on the RT site arrangements
The relative orientation of target sites with respect to one another determines the outcome of the recombination reaction: Cre or Flp will excise a circular molecule from between two directly oriented target sites or integrate a circular molecule into a linear molecule each possessing a target site **(A)**. A sequence flanked by a pair of RT sites with opposite orientation will be inverted **(B).** Translocation of DNA sequences occurs when a combination of two identical RT sites is present on two linear molecules **(C)**.

If recombinases are used to integrate transgenes into host cells, a stringent selection system is required, since the insertion reaction is thermodynamically unfavoured compared to the excision. A system that effectively meets this criterion is the so called "selection trap" which relies on the complementation of a non-functional, ATG-deficient resistance gene. The donor vector harbouring the gene of interest comprises a promoter and the missing start codon. Upon integration the resistance gene in the host cell genome is complemented and thereby transcriptionally activated. Consequently, resistance is exclusively conferred to correct recombinant daughter cells (O'Gorman *et al.*, 1991). This method was commercialised by Invitrogen as Flp-In System (Figure 1-8 A). However, this approach has some important limitations. Due to the reversibility of the reaction, recombinase activity has to be strictly limited. Furthermore prokaryotic sequences and a positive selection marker originally used to screen for successfully tagged cell lines are left behind in the chromosome after the integration (Wirth *et al.*, 2007).

A significant improvement was the development of non-interacting variants of the RT sites with mutated spacer sequences for both Flp and Cre target sites (Schlake & Bode, 1994; Bouhassira *et al.*, 1997). They can be used to flank the initial tagging cassette, thus allowing a precise exchange of the selectable marker with the sequences of an incoming donor vector harbouring the gene of interest. This method was termed recombinase mediated cassette exchange (RMCE). Its main advantage is the lack of excision and an increased frequency of targeting of up to 100 % due to stringent selection strategies (Schucht *et al.*, 2006). Additionally, no bacterial vector sequences are integrated which potentially limits mammalian gene expression (Figure 1-8 B).

RMCE has been demonstrated for Cre/loxP and Flp/FRT systems. Due to the thermolability of Flp at 37 °C Cre has been considered for a long time to be the most efficient system (Buchholz *et al.*, 1996). However, temperature optimised Flp mutants have been developed which show a 4-fold higher recombination efficiency (Buchholz *et al.*, 1998) and an increased activity in *in vitro* RMCE reactions (Takata *et al.*, 2011). Moreover the *de novo* synthesis of the mouse codon-optimised variant FlpO enabled an increase of recombination efficiencies to levels comparable to those of Cre (Raymond & Soriano, 2007). Thus, Flp mediated recombination is considered as superior compared to Cre/loxP system today, since several pseudo-loxP sites are present in the mammalian genome (Thyagarajan *et al.*, 2000). This technology has been successfully applied for the establishment of viral producer cells (Coroadinha *et al.*, 2006), for the evaluation of vector design (Verhoeyen *et al.*, 2001) and for the production of recombinant proteins such as EPO (Kim & Lee, 2008), or IgGs in CHO-K1 and HEK293 (Nehlsen *et al.*, 2009). At the Helmholtz Protein Sample Production Facility a RMCE master cell line based on the glycosylation deficient CHO Lec3.2.8.1 has been established especially for the production of glycoproteins for crystallization (Wilke *et al.*, 2011). Various proteins have been produced in our facility using this cell line by now, e. g. the lysosome-associated membrane glycoprotein 3 (LAMP3), the extracellular domain of the Toll-like receptor 2 (TLR2) and the γ-interferon-inducible lysosomal thioreductase (GILT) (Wilke, 2011) and numerous projects are still in the expression pipeline.

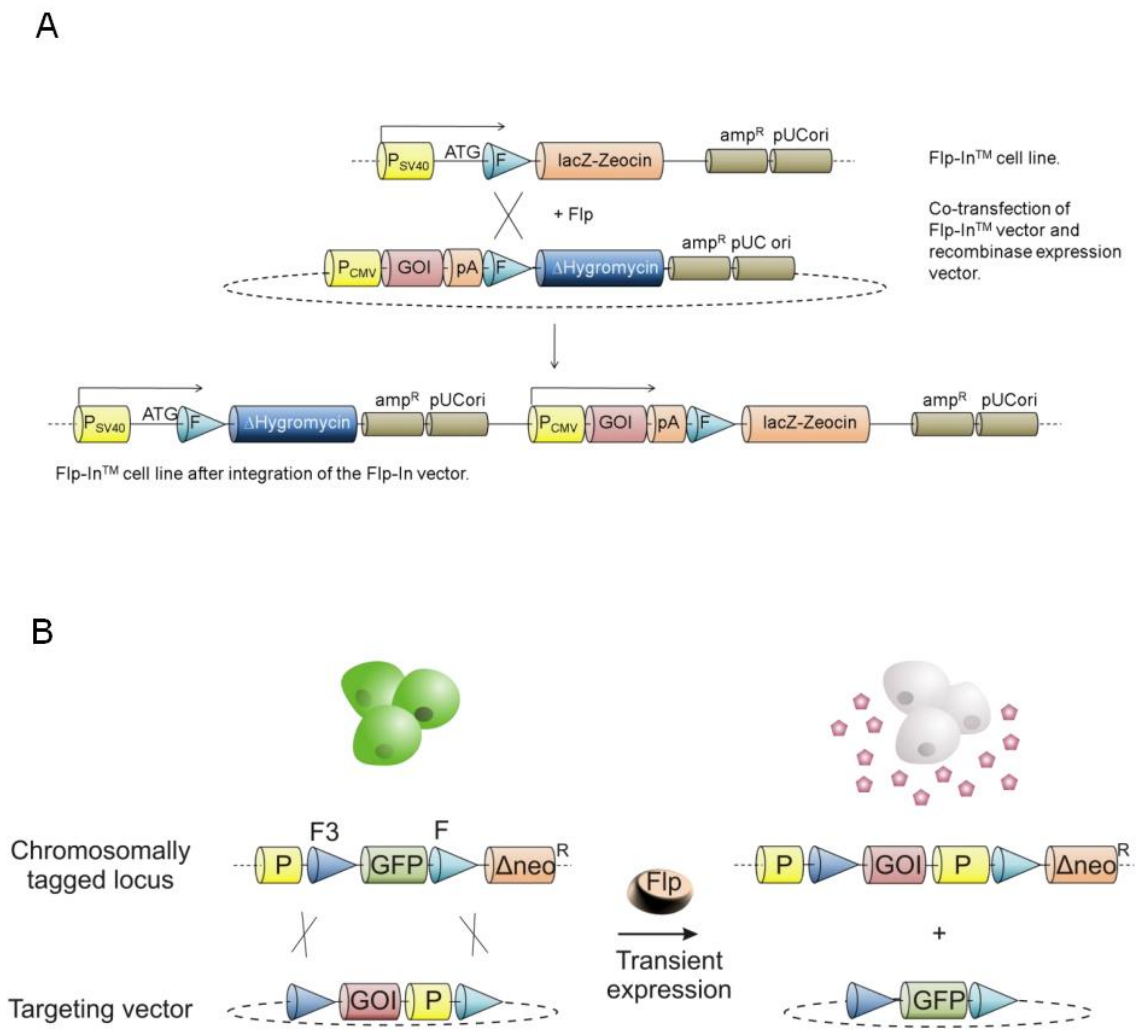


Figure 1-8: Schematic overview of the Flp-In™ system and RMCE
A: In the Flp-In™ system host cells are tagged at an expression hot spot by an expression cassette consisting of a simian virus 40 promoter (P_{SV40}), followed by an ATG start codon, one FRT site (F) and a *lacZ*-Zeocin fusion gene for selection. Upon cotransfection of the host cell line with a pFlp-In donor and a helper vector expressing the Flp recombinase, the integration of the donor is initiated. Those cells, where the GOI under the control of a pCMV is successful integrated can be selected due to the complementation of the ATG-deficient incoming hygromycin resistance gene. **B:** The RMCE master cell lines contain the GFP marker flanked by heterologous FRT mutants. Its expression is driven by a promoter located upstream of the leading F_3. A neomycin resistance gene missing its ATG start codon as well as a promoter (ΔneoR) lies downstream of the wild type FRT (F) site. The incoming donor vector harbours the gene of interest as well as a promoter followed by the start codon flanked by corresponding FRT sites in the same orientation. Upon RMCE, these elements are integrated upstream of the neomycin resistance gene which thereby becomes activated and confers resistance against G418. Cones = FRT sites (F_3 and F), P = promoter, pentagons = secreted protein product encoded by the GOI, ampR = bacterial beta-lactamase gene, pUCori = replication of origin, pA = poly adenylation site, GOI = gene of interest (source: Wilke, 2011).

1.5 Model proteins

The evaluation of promoter activities, optimisation of transfection methods or comparative analyses of expression systems are typically performed by the expression of model proteins. In order to be utilised as a reporter, a protein has to meet the following fundamental criteria: expressibility in the particular host, the absence or distinctiveness of endogenous homologous and the availability of sensitive as well as easy detection methods (Cullen, 2000). Commonly, enzymes such as chloramphenicol transferase (CAT) (Corsico & Howard, 1990), the firefly luciferase (Gould & Subramani, 1988) or secreted alkaline phosphatase (SEAP) (Berger *et al.*, 1988) are used, since well-established assays are available to detect and quantify their activity from cell lysates or culture supernatants, respectively. However, to achieve proof-of-concept studies with novel expression vectors as well as to perform quantitative and qualitative evaluations of different expression strategies and hosts, it is advisable to use proteins from different relevant protein classes and host species. In this work, test expressions to demonstrate the applicability of vector constructs and to determine expression characteristics in different hosts have therefore been performed with model proteins from three different protein classes. GFP variants of different size have been expressed as fluorescent reporters to proof intracellular accumulation of recombinant proteins. Production of secretory proteins has been evaluated using scFv-hFc fusions, since they represent a class of biotechnologically relevant proteins, which are usually produced with high yields. Additionally, expression characteristics of the extracellular domain of Toll-like receptors have been studied as an example of truncated, yet challenging protein fragments.

1.5.1 Fluorescent proteins

Expression of fluorescent reporter proteins can be directly detected by fluorescence microscopy or flow cytometry. Thus, no cell lysis, purification or assay-development is necessary. Additionally, an easy quantification is possible upon cell disruption via correlation with an fluorescent standard. The most prominent representative of fluorescent proteins is the green fluorescent protein, which originates from the jellyfish *Aequorea victoria* (Shimomura *et al.*, 1962). Because of its invaluable importance as a reporter in modern cell biology, its discoverer Osamu Shimomura has been awarded with the Nobel prize in 2008.

The 27.2 kDa large protein consists of 238 amino acids which form a β-barrel structure composed of 11 beta sheets as shown in Figure 1-9 (Ormo *et al.*, 1996). The amino acids Ser65, Tyr66 and Glu67 located in a α-helix in the interior of the barrel constitute the chromophore by autocatalytic formation of an aromatic system (Heim *et al.*, 1994). The excitation maximum of GFP is in the UV range at 488 nm and the emission peak is at 509 nm.

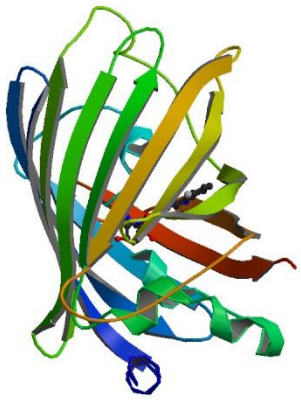


Figure 1-9: Structure of green fluorescent protein (GFP).
A β-barrel is formed by 11 β-sheets which surround a α-helix comprising the chromophore-forming amino acids S65, Y66, und E67. The molecule has a diameter of 2.4 nm and a length of 5 nm (source: PDB ID 1RM9).

Several variants of GFP have been developed by mutations of the amino acids forming the chromophore or in the surrounding environment, thereby creating molecules with different excitation and emission characteristics. The mutations S65T and F64L generated a significantly brighter variant by abetting the deprotonated form of the chromophore. Because of its optimised quantum yield and 35-fold higher brightness compared to wtGFP it was termed enhanced green fluorescent protein (eGFP) (Cormack *et al.*, 1996). Mutations of Y66 caused a shifting of the emission wavelength, thereby creating cyan and blue colour variations (CFP, BFP), respectively (Heim *et al.*, 1994; Heim & Tsien, 1996). Homologs of GFP from different species have also been identified and characterised. The red fluorescent protein drFP583 for instance, was isolated from the mushroom coral *Discosoma sp.* and is now commonly known as DsRed (Matz *et al.*, 1999). Its excitation peak is at 558 nm and a maximal emission can be measured in the red light range 583 nm. It is an obligate tetramer with the tendency to form oligomers, which can lead to protein aggregation. To optimise its expression qualities and fluorescent characteristics, mutations

within the chromophore and its vicinity have been performed, thus generating several variants with major differences in brightness, colour, stability and folding speed (Shaner *et al.*, 2004). Figure 1-10 shows a genealogic tree of available DsRed mutants.

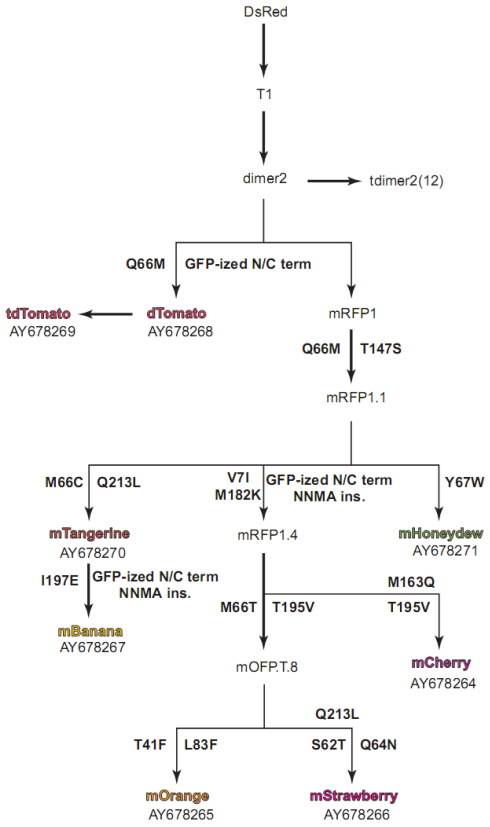


Figure 1-10: Genealogy of DsRed-derived variants
Amino acid mutations critical to the phenotype as well as insertions or exchange of N- and C-termini are annotated. GenBank accession numbers are given under the name of each protein. For detailed sequence alignments of the variants see original source (Shaner *et al.*, 2004).

Two DsRed variants have been used as fluorescent model proteins in this work. The first, tdTomato, is a covalently linked tandem dimer of dTomato which possesses the same emission maximum at 581 nm but improved folding characteristics compared to the physiological dTomato dimer. Upon expression, it forms a dimer of tandem dimers, resulting in a higher brightness. In contrary, mCherry, the second DsRed variant expressed as a reporter, is a strict monomeric mutant of DsRed with an emission maximum at 610 nm. It matures very rapidly ($t_{0.5}$ = 15 min) and is particularly stable against photobleaching. However, due to its monomeric form, it has only 30 % of the brightness of the DsRed tetramer (Shaner et al., 2004).

1.5.2 Single-chain antibody fragments

Recombinant antibodies are the fastest growing class of therapeutic proteins and exhibit an impressive track record. It took only one decade from the first generation of a monoclonal antibody in 1975 (Köhler & Milstein, 1975) to the approval of the first therapeutic antibody Murononab-CD3 which reduces acute rejection of transplanted organs (Goldstein *et al.*, 1985; Emmons & Hunsicker, 1987). After no significant increase in novel products until the year 2000, genetic engineering methods finally enabled two major steps which leveraged their extensive use for pharmaceutical applications: the generation of chimeric and humanised antibodies in the late 1980s (Jones *et al.*, 1986; Queen *et al.*, 1989) and the development of alternative antibody formats (Figure 1-11). Monovalent antibody fragments and engineered variants such as single-domain antibodies (also referred to as nanobodies) retain the targeting specificity of whole mAbs but can be produced more economically. Moreover, they possess unique properties thereby extending and improving the range of diagnostic and therapeutic applications. For instance, they show an improved tissue penetration, the ability to transmigrate the blood brain barrier, are suitable for *in vivo* detection or targeting of intracellular proteins (as so called intrabodies) and many more (Holliger & Hudson, 2005; Hagemeyer *et al.*, 2009; de Marco, 2011). Thus, today more than 400 clinical studies with antibody related therapeutics are ongoing (Dübel, 2007).

Besides their tremendous importance for therapeutical and diagnostical applications, antibodies are also used as tools for protein aggregation studies and crystallography chaperones (Hunte & Michel, 2002). Especially, single domain antibodies are used to improve crystal packing and X-ray phasing, as they give an ideal template for molecular replacement or can be used for SeMet labelling (Tereshko *et al.*, 2008). Since antibodies and antibody fragments are expressed and secreted with a high efficiency, impressive yields up to several g/L can be achieved in industrial processes (Durocher & Butler, 2009). Therefore, this protein class represents an excellent model for secreted proteins and to compare volumetric yields achieved in different expression systems and strategies.

In this work, scFv-hIgG1Fc-4E3, a CD30-specific single-chain Fv fused to a human Fc fragment (Menzel *et al.*, 2008) has been used for expression studies. Its antigen CD30, a 120 kDa transmembrane type I glycoprotein, belongs to the tumor necrosis factor receptor superfamily. The antibody was originally isolated from a semisynthetic human antibody

library by phage display and converted into the scFv-format at the TU Braunschweig. Likewise, a second variant of the antibody with a carboxy-terminal fusion of a human pancreatic RNAse (scFv-hIgG1-Fc4E3-RNase) was provided (Braschoss *et al.*, 2007).

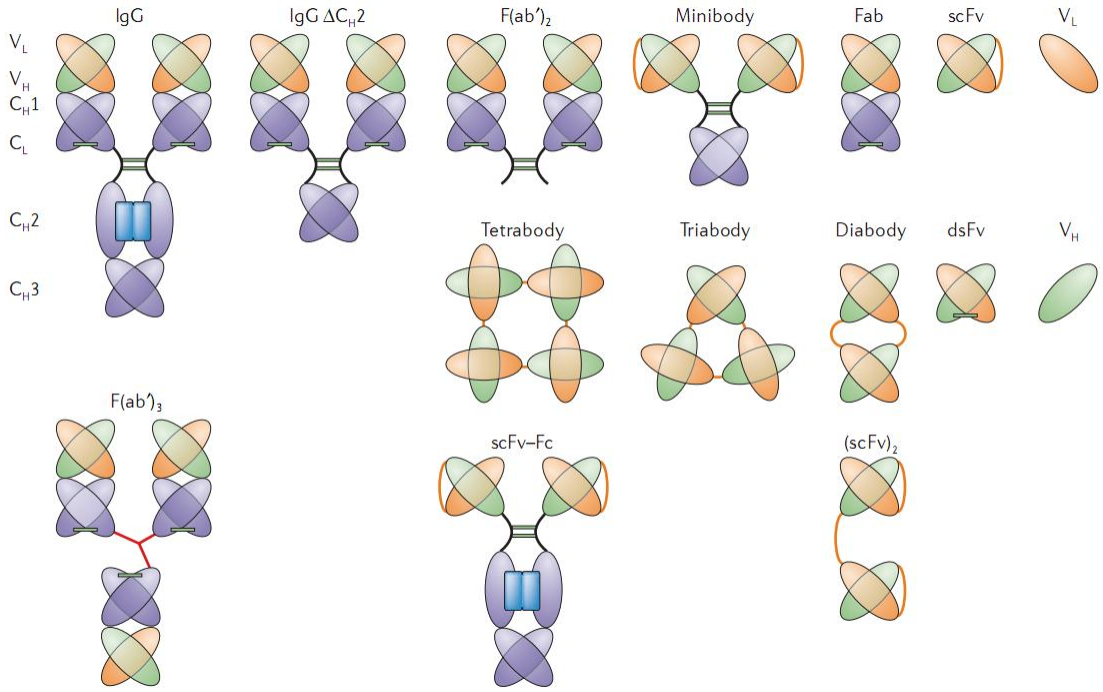


Figure 1-11: Immunoglobulins composed of combinations of heavy and light chains
IgGs are composed of two heavy and two light chains connected by disulphide bonds. Each chain consists of a variable (V) and one or more constant regions (C). The canonical antigen-binding fragments (Fab) are composed of variable heavy (VH), variable light (VL) and constant regions (Fc) of the light (CL) and heavy chains (CH). The building block that is most frequently used to create novel antibody formats is the single-chain variable (V)-domain antibody fragment (scFv), which comprises V domains from the heavy and light chain (VH and VL domain) joined by a peptide linker of up to ~15 amino-acid residues. The engineered fragments differ from the full-size IgG molecule in affinity, immunogenicity, and circulating half-life. Inter-chain disulphide bonds are depicted as green bars, glycosylation as blue boxes, peptide and chemical linkers are shown as orange and red lines, respectively. CH, C = domain of immunoglobulin heavy chain; CL, C = domain of immunoglobulin light chain; dsFv = disulphide-stabilised scFv (source: Carter *et al.*, 2006).

1.5.3 Toll-like receptors

Toll-like receptors belong to the family of the pattern recognition receptors (PRR), and are thus part of the innate immune system (Janeway & Medzhitov, 2002). The name was coined due to the similarity of TLRs to the drosophila protein Toll (Medzhitov *et al.*, 1997). These transmembrane receptors recognise molecules from exogenous origin via binding of conserved motifs, so called pathogen associated molecular patterns (PAMPs). Subsequently, a signal cascade is initiated, that finally results in the activation,

modulation and direction of immune responses (Sabroe *et al.*, 2008; Medzhitov, 2009). This principle is summarised in Figure 1-12 A. Until today, 13 TLRs have been found from which TLR1 - 9 seem to be closely homologous between mice and humans. In contrast, TLR10 is only found in the human genome, whereas TLR11, 12 and 13 are solely present in the mouse (Moresco *et al.*, 2011). According to the nature of their specific ligands, TLRs are located within the outer cell membrane or in intracellular compartments: TLR3, TLR7, TLR8, TLR9, and probably TLR13 of mice remain within the endoplasmic reticulum (ER), endosomes, multivesicular bodies, and lysosomes, whereas all other TLRs are secreted to the outer cell membrane (Blasius & Beutler, 2010). Similar to other immune receptor families, TLRs have a clearly defined domain organization, consisting of a ligand recognition domain, an effector domain and a transmembrane region which connects both domains. The ligand recognition domain comprises a series of leucine rich repeat (LRR) modules whereas the effector domain consist of a Toll/Interleukin-1 Receptor (TIR) motif which induces the downstream signalling (Figure 1-12 B).

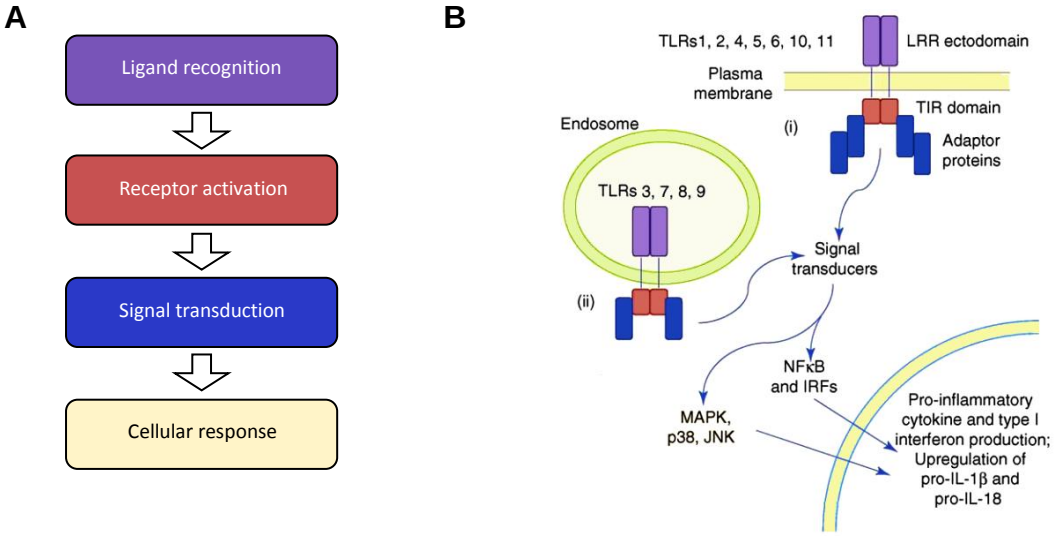


Figure 1-12: Simplified schematic of TLR signalling
TLRs sense ligands via their LRR ectodomains (purple) leading to receptor dimerization. Adaptor proteins (blue) are recruited to the dimeric TIR domain scaffold (red), creating a multiprotein complex located at either the plasma (i) or endosomal (ii) membrane. Adaptor protein usage varies with TLR. Signal transduction is propagated through the action of signal transducers such as the IRAK and tumor necrosis factor receptor-associated factor (TRAF) family members, as well as the mitogen-activated protein kinase (MAPK), p38 and c-Jun N-terminal associated kinase (JNK) pathways. Subsequently, the signalling leads to a cellular response that includes the production of pro-inflammatory cytokines and the upregulation of pro-IL-1β and pro-IL-18. This results in an increase in the levels of active transcriptional activators such as NFκB and members of the interferon response family (modified work, original source: Monie, *et al.* 2009).

The four adaptor proteins MyD88, TIRAP (also called MAL), TICAM1 (also called TRIF), and TICAM2 (also called TRAM) directly bind to the TIR domain of activated TLRs and recruit downstream signalling components (Monie *et al.*, 2009). They initiate a signal transduction that ultimately induces the expression of numerous genes required for the inflammatory response, such as inflammatory cytokines, chemokines antimicrobial molecules (e.g. hydrolytic enzymes, peptides, proteases), and major histocompatibility complex (MHC) as well as costimulatory molecules important for adaptive immune activation (Moresco *et al.*, 2011).

TLRs and other immune receptors not only play an important role in infection processes, but also contribute to chronic diseases or even induce potentially fatal clinical conditions such as sepsis upon inappropriate or excessive responses (Ospelt & Gay, 2010). Thus, understanding the mechanism of ligand-mediated receptor activation is considered as an important requirement for the development of improved therapies for multiple clinical conditions, antimicrobial treatments and vaccine adjuvants. Therefore, intense efforts were made to investigate crystal structures of the extra cellular domains (ECD) of TLRs in complex with their binding partners. In recent years, the structures of the extracellular domains of TLR1, 2, 3, 4 and 6 have been solved (Choe *et al.*, 2005; Jin *et al.*, 2007; Kim *et al.*, 2007; Kang *et al.*, 2009). Structure models of these TLRs in complex with their corresponding ligand are shown in Figure 1-13.

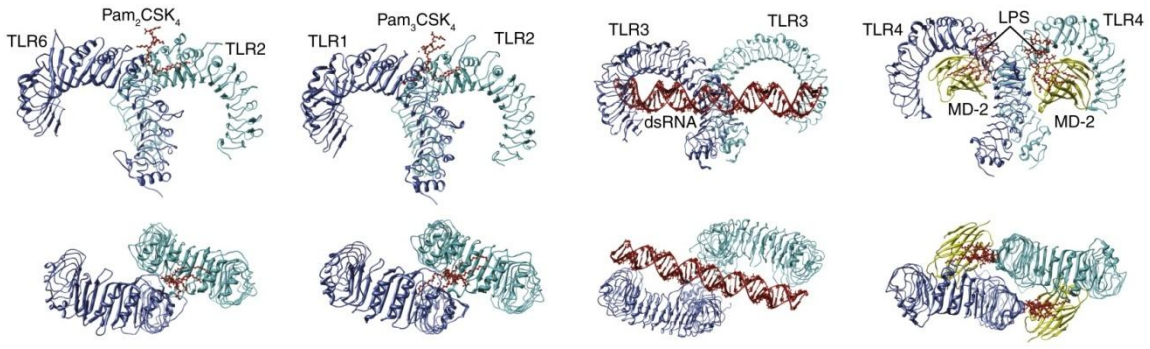


Figure 1-13: Structures of TLR dimers and interactions with ligands
Complex structures of TLR2–TLR6–Pam2CSK4 lipopeptide (3A79), TLR2–TLR1–Pam3CSK4 lipopeptide (2Z7X), TLR3-dsRNA (3CIY), and TLR4–MD-2–LPS (3FXI) are shown as side view (upper panels) and top view (lower panels). The Protein Databank ID numbers are indicated in parentheses (source: Moresco *et al.*, 2011).

The TLR-ECDs typically contain 19–25 consecutive LRRs that form a solenoid structure. Within each LRR, consensus hydrophobic residues point to the interior to build a stable core where three residues in the β-strand configuration align to form a hydrogen-bonded parallel β-sheet. Because the β-strands are more closely packed than the non-β portions of the LRR loops, the solenoid is forced into a curved horseshoe-like configuration (Kajava, 1998). As a result, each LRR protein contains a concave and a convex surface. The inner β-strands and the convex surface are connected via loops on the ascending lateral surface whereas a descending lateral surface is located on the opposite side (Bella *et al.*, 2008). In most LRR proteins, ligand binding occurs on the concave surface. By contrast, in the known TLR-ligand structures, ligand binding occurs most often on the ascending lateral surface of the TLR-ECD. The interaction is supported by the absence of N-linked glycans on this surface of the protein (Botos *et al.*, 2011).

Many LRR family proteins are difficult to produce in quantities sufficient for crystallization and purified material often could not be crystallised for X-ray diffraction studies (Jin & Lee, 2008). This fact underlines that a proper construct design with regard to domain boarders and secondary structure elements, is an essential prerequisite for soluble and stable protein products. In consequence, recombinant expression of the ECDs of human TLR1, 2, 4 and murine TLR2 and 6 for structure determination could only be achieved by means of an elaborate expression strategy called hybrid-LRR technology (Jin & Lee, 2008; Kang *et al.*, 2009). It relies on the creation of N- or C-terminal fusion constructs of TLR ECD truncations with fragments of the hagfish variable lymphocyte receptor 6 (VLRB6), an adaptive immune receptor discovered in jawless fishes, lamprey and hagfish (Pancer & Cooper, 2006). Since VLRs are also members of LRR protein family, it represents a structurally compatible fusion partner. The hybrid constructs have been generated by C- or N-terminal fusions of TLR and VLR fragments within the most conserved "LxxLxLxxNxL" motif. Since this sequence always forms β-strands upon expression, the secondary structure at the fusion position is conserved and a stable N- or C-terminus is generated by the corresponding termini of the VLR (Figure 1-14). However, despite the use of the hybrid-LRR technology only low volumetric yields of mTLR2 and 6 of not more than 100 µg purified protein from 1 L insect cell culture have been reported (Kang *et al.*, 2009). Recently, the full length ectodomain of hTLR4 fused to a Protein A tag could be coexpressed with MD-2, but likewise to the VLR hybrids of hTLR1, 2 and murine TLR4, the yields have not been disclosed (Park *et al.*, 2009).

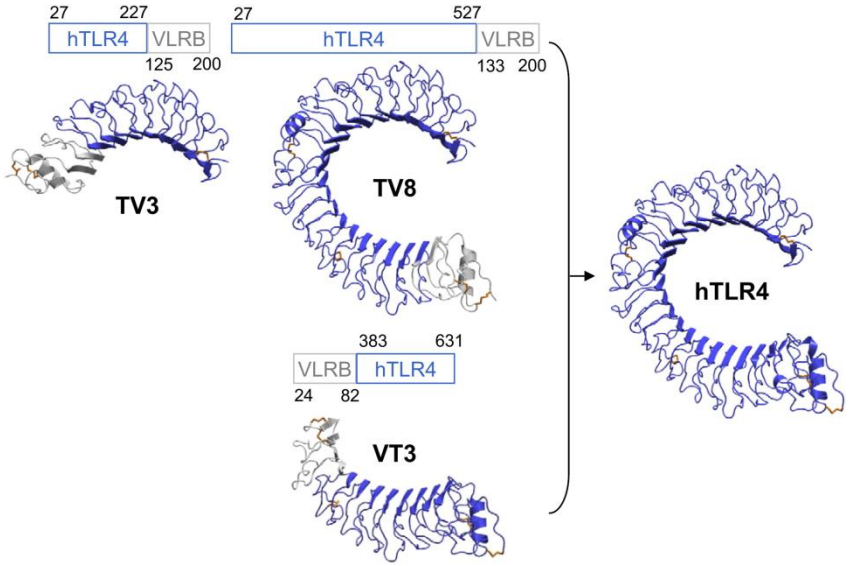


Figure 1-14: Hybrid LRR technology
Structures of the three human TLR4-VLRB.61 N-terminal (VT) and C-terminal (TC) VLR hybrids. The N-terminal fusions comprising amino acids 27 - 227 and 27 - 527 of hTLR4, whereas the C-terminal fusion comprises amino acids 383 – 631. The full-length ectodomain of human TLR4 was assembled after superimposition of the overlapping regions of the TV8 and VT3 hybrids. TLR4 fragments are coloured blue and VLRB.61 fragments gray. Disulphide bridges are represented in orange (source: Kim *et al.*, 2007).

Since Toll-like receptors are bridging innate and adaptive immunity (Pasare & Medzhitov, 2005) and thus are attractive targets for structural and biophysical investigations, several internal and external collaborative projects concerning TLRs are currently ongoing at the Helmholtz-Centre for Infection Research. This underlines the demand of improving recombinant expression of these targets. Therefore, the full length ectodomain of mTLR2 comprising amino acids 1 – 578 served as a model protein for difficult-to-express targets in this work. Additionally, the ECD of human TLR5, which structure was not known before, has been used for extended expressibility trials in different hosts. It represents an especially challenging target protein for recombinant production. Moreover, unlike the other TLRs, which structures have been solved, its main ligand is bacterial flagellin and thus a proteinaceous PAMP which makes it a particular worthwhile target for structural analyses.

1.6 Aim of this Work

Recombinant production of complex eukaryotic proteins for structure determination typically requires a profound screening process to identify expressible constructs. Furthermore, the evaluation of both the most suitable expression host and the optimal expression strategy has an important impact on protein quality, yield and cost efficiency. Hence, facilities dedicated to protein expression for structural biology like the Helmholtz Protein Sample Production Facility (PSPF) usually provide a broad spectrum of eukaryotic expression systems. Although vectors suitable for the use as expression- and donor plasmids in different systems have been published (Novy *et al.*, 1999; Berrow *et al.*, 2007), they suffer some major drawbacks, which limit their usability in multiparallel expression studies in state-of-the art systems. For instance, they are not compatible to advanced transposition based techniques for the generation of recombinant bacmids and lack the EBVoriP for enhanced expression in optimised HEK derived cell lines. Additionally, they are not applicable for stable genomic expression in mammalian cells by the Flp-recombinase mediated cassette exchange system (RMCE).

The primary aim of this work was therefore the development of an improved expression/donor vector for protein production in the PSPF. Its range of applications should comprise different expression methods including transient, viral and stable expression strategies to facilitate both a rapid screening for expressible protein variants as well as the evaluation of the optimal expression host for any target protein. Thereby, the vector should enable the direct shifting to large-scale production with the proper construct in the system of choice avoiding time-consuming recloning steps. Subsequent to the construction of this vector, a proof-of-concept should be provided by test expression of a set of different model proteins to demonstrate its qualification.

The second part of the project aimed at the implementation of the RMCE system to insect cells, allowing the generation of specialised BEVS host cell lines for stable expression of auxiliary proteins to aid or enhance baculoviral expression of challenging target proteins. By the combination of a new versatile vector for different expression systems and the implementation of a compatible RMCE platform for mammalian and insect host cell lines, an integrated expression system at the PSPF should be provided, to facilitate an efficient multiparallel screening and thereby streamline the expression pipeline.

2 Material and Methods

2.1 Chemicals, kits and reagents

If not stated otherwise, all chemicals used in this work were of "pro analysis" grade and were purchased from the companies Amersham Biosciences, Difco, Fluka, GE Healthcare, Gibco, Invitrogen, Macherey-Nagel, Merck, Millipore, QIAGEN, Riedel de Haen, Roche, Roth, Sigma and Stratagene.

2.1.1 Enzymes and molecular weight standards

Commonly used enzymes and molecular weight standards are listed in Table 2-1 and Table 2-2.

Table 2-1: Enzymes

Name	Supplier
Restriction Endonucleases	New England Biolabs
T4 Ligase	Roche
Antarctic Phosphatase	New England Biolabs
SAP (Shrimp Alkaline Phosphatase)	Roche
Platinum Pfx DNA Polymerase	Invitrogen
Phusion HotStart DNA Polymerase	Finnzymes
DNAse	MOSB, HZI

Table 2-2: Molecular weight standards

Name	Usage	Supplier
Smart Ladder	agarose gel electrophoresis	Eurogentech
PageRuler Plus prestained	SDS-PAGE	Fermentas
PageRuler unstained	SDS-PAGE	Fermentas
Precision Plus unstained	SDS-PAGE	BioRad
Precision Plus AllBlue	SDS-PAGE	BioRad

2.1.2 Culture media & supplements

Cultivation of bacterial cultures was performed using LB medium (Bertani, 1951) supplemented with appropriate antibiotics or reagents for blue-white selection (Table 2-3). For resuspension of transformed strains subsequent to electroporation SOC medium (super optimal broth for catabolite repression) was used (Table 2-4). Media for bacterial cultures were autoclaved 20 min at 121 °C. Heat instable components such as glucose and antibiotics were added after filtration (0.2 µm).

Table 2-3: Antibiotics for bacterial cultures

Antibiotics	Final concentration
Ampicillin (Amp)	100 µg/mL
Chloramphenicol (Cm)	25 µg/mL
Gentamicin (Gm)	7 µg/mL
Kanamycin (Kan)	50 µg/mL
Tetracyclin (Tet)	10 µg/mL
Bluo-Gal	100 µg/mL
IPTG	40 µg/mL

Table 2-4: SOC medium

Composition	Final concentration
Bacto-Tryptone	20 g/L
Yeast extract	5 g/L
KCl	0,186 g/L
$MgCl_2$	10 mM
NaCl	50 mg/L
Glucose	10 mM
in ddH_2O	

Cell culture media for insect and mammalian cell lines (Table 2-5) were purchased as ready-to-use solutions or prepared from powdered formulations. Powder media was sterilised after dissolving in purified water (MilliQ) by a filtration line (Pall) combining two nylon filter units with decreasing pore size (1 µm, 0.2 µm). Supplements were added according to Table 2-6 after filtration (0.2 µm)

Table 2-5: Cell culture media

Cell culture medium	Cell lines	Supplier
EX-CELL™ 405	Hi5	SAFC
EX-CELL™ 420	Sf21	SAFC
F17	HEK293-6E	Gibco
ProCHO5	SWI3-26 (CHO Lec3.2.8.1)	Lonza

Table 2-6: Supplements

Supplements used in F17 Medium	Final concentration
G418	0,025 mg/L
Pluronic F68	1 g/L
L-Glutamine	7,5 mM

Supplements used in ProCho5	Final concentration
Phenolred	11 g/L
L-Glutamine	7,5 mM

Supplements	Supplier
Fetal Calf Serum (FCS)	Gibco
Fungizon (250 µg/mL)	Gibco
Gentamicin (10 mg/mL)	Gibco
Geneticin G418 (50 µg/mL)	Gibco
Penicillin (5000 U/mL) / Streptomycin (5 mg/ml)	Gibco
Pluronic F68 (100 g/L)	Invitrogen
L-Glutamine (200 mM)	Invitrogen
Phenol red	Sigma
Tryptone N1, cat # 19553	Organotechnie S.A.S.
Valproic acid (80 mM)	Sigma-Aldrich

2.1.3 Transfection reagents

Delivery of recombinant DNA to cells was performed using cationic lipofection techniques. The reagents used for each cell line and DNA construct as listed in Table 2-7.

Table 2-7: Transfection reagents

Reagent	Use	Supplier
Cellfectine	Transfection of insect cells	Invitrogen
Insect Gene Juice	Transfection of insect cells	Novagen
Lipofectine	Transfection of insect cells	Invitrogen
Superfect	Bacmid Transfection	QIAGEN
Polyethylenimine (linear, MW ~25 kDa)	Transfection of HEK293-6E	Polysciences
MetafectenePRO	Transfection of CHO	Biontex

2.2 Oligonucleotides and plasmids

Oligonucleotides were used for sequencing, site directed mutagenesis and cloning of DNA-fragments. They have been ordered at MWG Eurofins Operon in HPLC purified quality. A complete list of all used primers is in the appendix.

All plasmids generated in this work or used as parental vectors for the construction of pFlpBtM variants, tagging vectors and donor vectors of model proteins respectively are listed below.

pFlpBtM variants

pFlpBtM-I
Vector for transient expression in insect cells and donor for Tn7 transposition based generation of recombinant bacmids. It also contains FRT sites compatible for RMCE in the CHO Lec3.2.8.1 master cell line SWI3-26. Construction, gene integration and evaluation is presented and discussed in this work.

pFlpBtM-II
Multi host expression and donor vector for transient expression in mammalian cell lines and Tn7 transposition based generation of recombinant bacmids. It also contains FRT sites compatible for RMCE in the CHO Lec3.2.8.1 master cell line SWI3-26 and an inducible T7 promoter for expression in *E. coli*. The version pFlpBtM-II(beta) has no Shine-Dalgarno sequence and differs in the design of the MCS and included tags. Furthermore it lacks restriction sites flanking FRT sites and promoters. Construction of pFlpBtM-II variants, gene integration and evaluation is presented and discussed in this work.

Tagging vectors

pie1-eGFP-dneo
Tagging vector used for the integration of an RMCE cassette with neomycin selection trap in Sf21 and Hi5. It is a derivative of pFS-eGFP-dneo (Wilke, 2011) where the EF-promoter was excised by BglII and HindII and replaced by a PCR-fragment containing the ie1 and p10 promoter of pIEX/Bac5.

pie1-tdTomato-dneo
Derivative of pie1-eGFP-dneo. The eGFP sequence was excised by BstBI and BsaBI and replaced by a PCR-fragment containing tdTomato.

Vectors used for construction

pFastBac (Invitrogen)
Standard donor vector of the Bac-to-Bac™ system. This plasmid was the basis for the construction of pFlpBtM vectors.

pIEx/Bac 5 (Novagen)
Dual purpose vector for transient and viral expression in insect cells. It contains a promoter fragment with the baculoviral immediate early 1 and the late viral p10 promoter. This plasmid was used as a template for generating a PCR-fragment of the promoter region to construct pFlpBtM-I and pie1-eGFP-dneo.

pTriEx (Novagen)
Expression vector for transient expression in vertebrate cells, in *E. coli* and donor vector for homologuos recombination with triple cut baculoviral DNA to generate recombinant baculoviruses. It was used as a template to generate a PCR-fragment of the promoter region containing a CMV ie promoter/enhancer, the baculoviral p10 promoter and a T7/lac sequence for the construction of pFlpBtM-II

pTT5 (NRCC)
Expression vector for transient expression in HEK293 cells. It contains the Epstein-Barr virus oriP sequence for episomal replication and enhanced nuclear transport in cells expressing the Epstein-Barr nuclear antigen 1 (EBNA1). This vector served as a reference in transient expressions in HEK293-6E. Furthermore, it was the template for the generation of a PCR-fragment containing the EBNA1 oriP, the beta-lactamase gene and the pMB-ori upon deletion of NcoI and BbsI sites in the backbone by site-directed mutagenesis. This fragment was used for the construction of pFlpBtM-II.

pTT5-VLR-recipient
Derivative of pTT5 with an integration of a hagfish variable lymphocyte receptor sequence followed by a TEV protease site and a one-STrEP- and His_8 tag. It was used as a template to generate a PCR-fragment containing the TEV-STrEPHis8 sequence which was subsequently integrated into pFlpBtM-II

pFS-eGFP-dneo (S. Wilke, HZI)
Tagging vector used for the integration of an RMCE cassette with neomycin selection trap in CHO Lec3.2.8.1 (Wilke et al., 2011). A promoterless neomycin phosphotransferase (Δneo) gene lacking the ATG codon is located downstream of GFP. When the reporter is exchanged by a cassette including a promoter and start codon, Δneo becomes complemented. Thus, the cell gains resistance against the antibiotic G418 allowing for selection of recombinant cells.

pFS-sighis-PGKd (S. Wilke, HZI)
Targeting vector to integrate genes into the RMCE locus. It is compatible with the FRT sites of pFSeGFP-dneo and contains a PGK promoter with a ATG codon downstream of its MCS for the selection trap system. This vector was used as a PCR-template to amplify a fragment consisting of the FRT sites, the MCS and the elements for the selection trap. This fragment was integrated into pFlpBtM-I.

Parental vectors with inserts

pFBDM-mTLR2 (N. Kuklik, HZI)
pFBDM-mTLR2 is a donor vector of the MultiBac system to generate recombinant bacmids. It contains sequence of the ECD of murine TLR2 (uniprot accession no. Q9QUN7, kindly provided by C. Kirschner, Univ. Essen). The plasmid was used as a PCR-template to generate a fragment to be integrated into pFlpBtM vectors as a model protein.

pmCherry & ptdTomato (Clontech)
These vectors contain the sequences of the DsRed mutants mCherry and tdTomato, respectively (Shaner et al., 2004). They have been purchased from Clontech and were used as a PCR-template to generate fragments of each gene to be integrated into pFlpBtM vectors as model proteins.

pCMV-scFv-hIgG1Fc-4E3 (C. Menzel, TU Braunschweig)
Mammalian expression vector harbouring the sequence of an scFv-hIgG1-Fc fusion construct (Menzel et al., 2008). It was used as a reference in transient expressions in HEK293-6E and served as a PCR-template to generate a fragment of the scFv-hIgG1Fc gene which was cloned into pFlpBtM-II and pTT5.

pTT0/GFPq (NRCC)
This vector is a variant of pTT and harbours an GFPq-gene (red shifted green fluorescent protein). It is used as a transfection control plasmid in transient expression in HEK293-6E cells. Transfection efficiency can be determined by flow cytometric measurements of fluorescent cells upon transfection.

2.3 Bacterial strains and cell lines

2.3.1 Bacterial strains

Cloning steps to construct new plasmids or integrate coding sequences into expression vectors were performed in Xl1 Blue or Top10 *E. coli* strains. For the generation of recombinant baculoviral DNA DH10 strains harbouring the corresponding bacmids were used. All strains are listed in Table 2-8.

Table 2-8: Bacterial strains

***E. coli* strain**	**Genotype**	**Source**
Top10	F-, *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80(*lac*Z) ΔM15Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara, leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
XL1-Blue	endA1 gyrA96(nalR) thi-1 *rec*A1 *rel*A1 *lac* *gln*V44 F'[::Tn10 proAB+ *lac*Iq Δ(*lac*Z)M15] *hsd*R17(r_K^- m_K^+)	Stratagene
DH10MultiBac DH10EMBacY	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80(*lacZ*) ΔM15ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 *gal*U *gal*K l- *rps*L *nup*G /BacloxP/ pBADZ-His6Cre/pMON7124	EMBL (Berger)

2.3.2 Cell lines

The following mammalian and lepidopteran cell lines have been used as host cell lines for protein expression in this work.

CHO Lec3.2.8.1

CHO Lec3.2.8.1 is a glycosylation mutant clone derived from the parental chinese hamster ovary (CHO) clone Pro-5 (a proline auxotroph, ATCC no. CRL 1781) by selection for resistance to different lectins (Stanley, 1989). This clone was adapted to grow in suspension.

HEK293-6E

The human embryonic kidney epithelial cell line HEK293 was established in 1977 by transformation of a primary culture with a adenovirus (Graham *et al.*, 1977). The subclone HEK293-6E (NRC Canada) is adapted to suspension and carries an integration of a truncated Epstein-Barr nuclear antigen 1 (EBNA1) (Durocher *et al.*, 2002).

IPLB-SF-21AE (Sf21)

The insect cell line IPLB-SF-21AE (DSMZ no. ACC 119) was originally isolated from ovaries of the fall armyworm *Spodoptera frugiperda* (Vaughn *et al.*, 1977). It is adapted for growth in suspension and used in the BEVS for primary infection, virus amplification and protein production.

BTI-Tn-5B1-4 (Hi5)

BTI-Tn-5B1-4 (Wickham & Nemerow, 1993) is an ovary cell line which originates from the cabbage looper *Trichoplusia ni.* It is a subclone of the cell line BTI-Tn-5B1-28 (Granados *et al.*, 1986) and is commercialised under the brand name High Five™ by Invitrogen. Due to higher expression rates this cell line is an alternative host for protein production using recombinant baculoviruses.

2.4 Molecular biological methods

Molecular-biological methods used in this work comprise plasmid preparation, PCR, restriction digest, extraction of DNA from agarose gels, ligation of DNA fragments, site directed mutagenesis and transformation. The corresponding protocols have been adapted from standard collections of methods and protocols (Sambrook & Russel, 2000; Asubel *et al.*, 2007) and will not be explained here in detail. Isolation of recombinant bacmid DNA was performed according to the MultiBac manual (Berger *et al.*, 2008). Subsequent to cloning procedures, plasmid constructs were sequenced at the Genome analysis department at the HZI.

2.5 Cell culture techniques

2.5.1 Maintaining cells in culture

If not stated otherwise cell lines were maintained as suspension cultures in 125 mL (culture volume 20 – 50 mL) or 500 mL shake flasks (culture volume 100 – 200 mL) on orbital shakers (Infors TR250, Biometra WT12 or Kühner ES X). Media and supplements for each cell line were used according to Table 2-5 and Table 2-6 on page 34.

HEK293-6E and CHO Lec3.2.8.1 were cultivated at 37 °C and 100 rpm in humidified atmosphere with 5 % CO_2. Adherent cultures were incubated at 8 % CO_2. To maintain cells in suspension in the exponential growth phase, they were subcultured in fresh medium every 3 to 4 days with a seeding cell density of 1.5-3 x 10^5 cells/mL. Adherent cultures were expanded to larger vessels or transferred to suspension culture when reaching confluency.

Sf21 and Hi5 were cultivated at 27 °C and passaged every 2 to 3 days to a cell density of 3-7 x 10^5 cells/mL. Suspension cultures have not been humidified or maintained in a gassed atmosphere, whereas adherent cultures were stored in a humidified incubator to prevent evaporation.

2.5.2 Determination of cell number and viability

Cell numbers and viability were assayed by trypan blue dye exclusion method in a Neubauer hemocytometer or by flow cytometry after staining with propidium iodide. Both dyes cannot permeate intact cell membranes so that only dead cells are stained. Trypan blue binds to proteins whereas propidium iodide intercalates with nucleic acids. For manual determination of the cell density and viability 25 µL cell culture were diluted in 75 µL 0.5 % trypane blue solution. Both stained and unstained cells of 4 big squares of the hemocytometer were counted and the cell density and viability were calculated according to the following equations:

$$\frac{\textbf{cells}}{\textbf{mL}} = \frac{\textbf{cells} \times \mathbf{10^4} \times \textbf{dilution factor}}{\textbf{no.of big squares}} \qquad \textbf{Eq. 1}$$

$$\textbf{viability}\ [\%] = \frac{\textbf{viable cells} \times \mathbf{100}}{\textbf{total cells}} \qquad \textbf{Eq. 2}$$

The determination of cell density and viability by flow cytometry was performed using a Guava EasyCyte™ Mini System (Millipore). Cell suspensions were diluted 10-fold in PBS and stained with a final concentration of 50 µg/mL propidium iodide to identify dead cells. The principle of flow cytometry is described in detail in 2.5.6 on page 42.

2.5.3 Cell size determination

Since an increase in insect cell diameter of 10 % to 20 % can be observed upon baculoviral infection, measurements of cell size variations have been performed in this work to monitor the progress of the infection. The CASY Cell Counter (Innovatis) facilitates a simultaneous measurement of cell number, viability and size due to its high acquisition rate of 10^6 min^{-1}. The device utilises the principle of electric current exclusion for the determination of the size of particles diluted in an electrolyte. While flowing through a capillary the particles pass an electrical field and induce a resistance which is proportional to their size. Since dead cells exhibit membrane permeabilization, the resistance induced by those cells is significantly lower. Thus, a determination of the viability of a cell suspension is possible following a cell type specific calibration. The measurements were performed according to the manufactures guidelines after the samples were diluted in an isotonic, particle pure buffer to reach the measurement range of the device.

2.5.4 Cryopreservation

Cell banks of cell lines were stored in the vapour phase over liquid nitrogen. To prepare the cells for long term storage they were harvested from suspension culture by centrifugation at 180 ×g for 4 min. The cell pellet was resuspended in freezing medium (45 % (v/v) culture medium, 45 % fresh medium and 10 % DMSO) at a cell density 2×10^7 cells/mL. Each cryo-tube (Nunc) was filled with 1.8 mL cell suspension. To provide a constant cooling rate of -1 °C/min the vials were stored in a special freezing container (Nalgene) in isopropanol at -70 °C for 24 h. Subsequently, the cryo-tubes were transferred into the liquid nitrogen container for long term storage.

2.5.5 Chemical based transfection of eukaryotic cells

Transfection of cell lines with plasmids or recombinant viral DNA (i.e. bacmids) was performed by cationic lipofection techniques. This process relies on complex formation of negatively charged DNA with lipofection reagents that possess linear or branched structures terminating at positively charged amino groups. Since the DNA-reagent

complexes show a net positive charge, they bind to negatively charged receptors (e.g., sialylated glycoproteins) on the surface of eukaryotic cells. The compact structures optimise the entry of DNA into the cell. Upon endocytosis and fusion with the endosome, the reagents are capable of pH-buffering the lysosome ensuring inhibition of lysosomal nucleases and stability of the DNA complexes.

Different lipofection reagents are composed of different cationic or neutral polymers (or different ratios thereof) such as N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA) (Felgner *et al.*, 1987), dioleoyl phosphatidylethanolamine (DOPE) (Felgner *et al.*, 1994) or N,N^{I},N^{II},N^{III}-tetramethyl-N,N^{I},N^{II},N^{III}-tetrapalmityl-spermine (TM-TPS) (Gao & Huang, 1995). Hence, for the transfection of CHO and insect cells optimised transfection reagents and conditions for each DNA construct have been evaluated according to the manufacturer's instructions. HEK293-6E cells were transfected using linear polyethylenimine (PEI) with an average MW of 25 kDa. The polycationic PEI is the most efficient reagent for the delivery DNA into HEK293 cell lines (Huh *et al.*, 2007). Additionally, DNA-PEI aggregates protect the nucleic acids from degradation by nucleases (Boussif *et al.*, 1995). The protocols used in this work for the transfection of insect and mammalian cells are described below. The corresponding reagents are listed in Table 2-7 on page 34.

Transfection of insect cells with plasmid DNA

Insect cells in serum free suspension cultures were transfected with plasmid DNA for transient tests expression or the genomic integration of RMCE cassettes. For this purpose Lipofectin, Cellfectin (both Invitrogen) and Insect Gene Juice (Novagen) have been used in DNA to reagent ratios of 1:2 - 1:4 with a DNA dose of 1 - 2 µg/10^6 cells. Complex formation was performed according to the manufacturer's protocols: after dilution of DNA and reagent in appropriate amounts of medium the solutions were mixed and incubated at RT for 10 - 20 min to allow complex formation. A removal of transfection complexes after adding to the cell culture by a medium exchange was implemented if recommended.

Transfection of insect cells with bacmids

To generate recombinant viruses adherent Sf21 cells were transfected with recombinant bacmids of the MultiBac-system or EMBacY-bacmids, respectively. For transfection 0.75×10^6 Sf21 cells from a culture supplemented with 5 %FCS per well were seeded on 6-well plates. For each transfection 10 µL Superfect (QIAGEN) and 5 µL isolated bacmid were diluted in a total volume of 200 µl and incubated for 20 min at RT prior to adding the mixture to the cells. Whereas the formation of transfection complexes was performed under strict absence of serum, the incubation of the cells with the complexes was performed in serum supplemented medium. After 2 h the transfection mixture exchanged by 2 mL fresh medium. The supernatant containing first generation virus was harvested 72 h post transfection.

Transfection of HEK293-6E

Transfection of HEK293-6E was performed using optimised protocols provided by the NRCC, with a DNA dose of 0.2 µg per 10^6 cells and a DNA:PEI ratio of 1:2. 25 µg plasmid DNA and 50 µg PEI were each diluted in 1.25 mL supplemented F17. After both solutions were mixed and incubated for 15 min at RT to allow complex formation, the mixture was quantitatively added to exponentially growing cells in 25 mL suspension culture with a density of 1.5 - 2 × 10^6 cells/mL. Prior to complex formation the DNA was spiked with 5 % pTT0/GFP, allowing to monitor the transfection efficiency by flow cytometric measurements.

2.5.6 Flow cytometry and preparative cell sorting

Fluorescence based flow cytometry is a method for real time high-throughput analysis of cells in suspension. Cells are diluted in a hydrodynamically focussed sheath fluid and excited by a single wavelength laser which allows a simultaneous multimeric detection of particle size, granularity and fluorescence. It was evolved in 1972 (Julius *et al.*, 1972) from a previously presented method which is based on the Coulter principle of cell size determination to separate cells (Fulwyler, 1965). In the prevalent optical based method the cell size is measured by a detector in the line of the beam (forward scatter), whereas granularity (side scatter) and fluorescence is simultaneously determined by perpendicular detectors (Figure 2-1, upper part). The data is analysed by a software and presented as histograms or two-dimensional scatter plots. Clusters of cells with similar properties can be identified which enables the definition of so called gates to mark certain subpopulations for evaluation purposes. Thereby, a determination of transfection efficiencies or viability is possible on the basis of the number of eGFP-positive cells or cells stained by propidium iodide, respectively.

Besides analytic flow cytometers so called cell sorters can be used for preparative fluorescence activated cell sorting (FACS) according to measured properties of the cells. In this method cells belonging to certain subpopulations marked by gates in the evaluation software will be separated from the suspension during the sorting process. Therefore the liquid column of the cell solution is separated into single droplets by a vibrating piezoelectric crystal. It causes velocity fluctuations in the emerging liquid column which leads to growing disturbances of the surface tension. Immediately after the jet passes the laser beam it breaks into very uniform droplets each containing a single cell. These droplets are subsequently equipped with a positive or negative charge according to the gate the enclosed cell has been assigned to by the evaluation. Following to droplet formation, the cells pass through an electrostatic field, which causes deflection of charged

droplets which are collected in separate containers while uncharged droplets remain undeflected and are discarded. The automated sorting process includes the measurement of cell properties, data evaluation, signal feedback, separation of empty droplets or those containing cell aggregates as well as the sorting of positive cells in up to four different populations. State-of-the-art devices achieve typical acquisition rates of 25,000-70,000 events per second, depending on the demands in purity and yield.

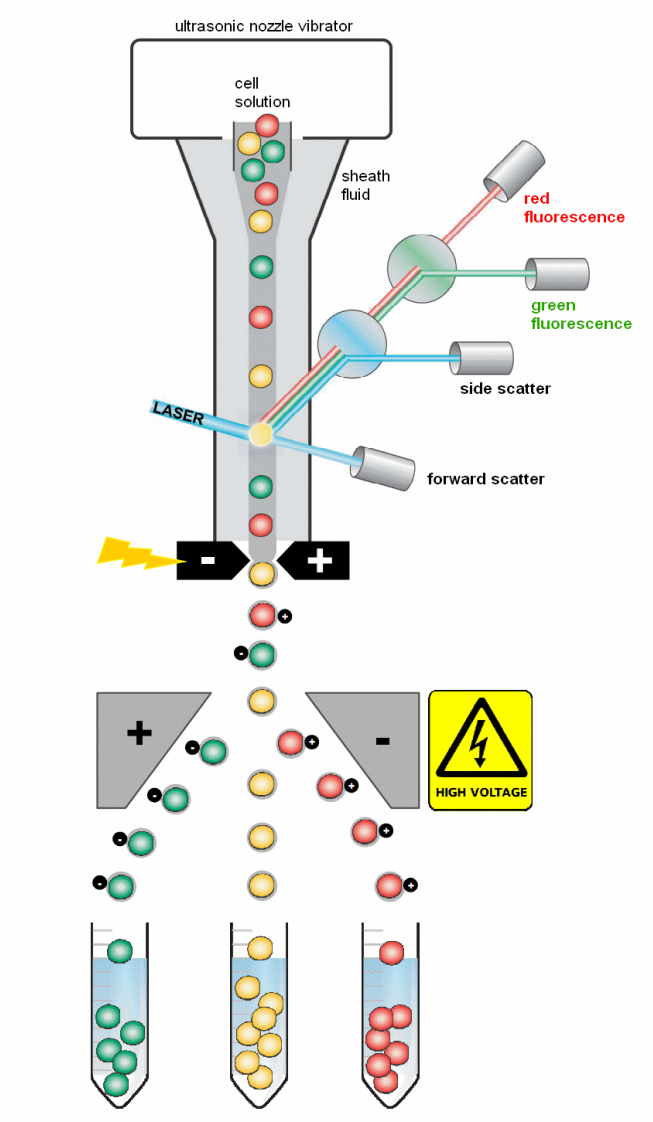


Figure 2-1: Principle of flow cytometry and cell sorting
Cells diluted in a sheath fluid pass a laser beam which allows a simultaneous detection of cell size (forward scatter), granularity (side scatter) and fluorescence by photomultipliers in the line of the beam and perpendicular to it, respectively. A piezoelectric element causes the jet to separate in droplets containing single cells immediately behind the laser beam. The droplets are electrically charged after passing the nozzle according to the data evaluation feedback. While passing through an electrostatic field, the charged droplets are deflected and thereby can be sorted into up to four different pools (modified work, original source: L. Gröbe, HZI).

In this work a Guava EasyCyte Mini System (Millipore) has been used for the determination of cell number, viability and efficiencies of both transfections and infections with recombinant baculoviruses. It is equipped with a 488 nm argon laser and 3 band-pass filters for the detection of green, yellow and red emission signals. Preparative cell sorting was performed on a MoFlo high-speed cell sorter (Beckman Coulter) and a FACS Aria II™ cell sorter (Becton Dickinson) by Lothar Gröbe in the research group Experimental Immunology at the HZI. The FACS Aria II was equipped with two frequency-doubled diode-pumped solid-state laser with wavelengths of 488 nm and additionally 561 nm for the excitation of tdTomato. For the isolation of single cells an automated cell deposition unit sorting cells into a 96-well plate was used at the MoFlo system. Table 2-9 shows the configurations of each device used in this work.

Determination of fluorescence using the Guava EasyCyte Mini was performed according to the protocol described in 2.5.2 without staining with propidium iodide. Preparative cell sorting was performed with aliquots of at least 3×10^7 cells, which were centrifugated at 180 ×g for 4 min. The supernatant was sterilised by filtration and collected as conditioned medium. The cells were resuspended in 2 mL fresh medium supplemented with 10 % FCS and Penicillin, Streptomycin and Fungizon. Sorting was performed until 3×10^7 positive cells were collected in 5 mL medium supplemented with FCS and antibiotics. These cells were pelleted and resuspended in 2 mL conditioned medium suplemetented with FCS and antibiotics and seeded in a single well of a 6-well plate. After 24 hours medium was exchanged by conditioned medium. The cells were proliferated until confluency and passaged or expanded into 4 mL suspension culture.

Table 2-9: Configuration parameters of flow cytometer

Sorting device	Marker protein	Excitation laser	Detector
Guava EasyCyte Mini	eGFP / YFP	488 nm	525/30
	tdTomato	488 nm	583/26
	mCherry	488 nm	680/30
BD FACS Aria II	eGFP	488 nm	525/50
	tdTomato	561 nm	585/15
Beckman MoFlo	eGFP	488 nm	530/40
	tdTomato	488 nm	580/30

2.5.7 Generation of producer cell lines by RMCE

The delivery of RMCE donor-plasmids and the helper-plasmid harbouring the Flp recombinase into the CHO Lec3.2.8.1 derived RMCE master cell line SWI3-26 was performed by nucleofection. This is an electroporation-based method to deliver substrates into the cytoplasm and the nucleus of mammalian cells. Since it is a direct transfer this technique is independent of cell division for the transport of DNA into the nucleus. A proprietary variant commercialised by Lonza relies on a combination of cell line specific electroporation parameters and optimised reagents for different cell lines. The cotransfection of the master cell line and the subsequent isolation of producer cell lines was performed by Bahar Baser using the following protocol.

Nucleofection of CHO Lec3.2.8.1

1×10^6 cells were resuspended in 100 µL of nucleofector-solution transfected with 1 µg donor- and 4 µg helper plasmid by using program U-24 according to manufacturer's guidelines (Nucleofector Kit V, Lonza). Directly after electroporation the cells were resuspended in 500 µL prewarmed supplemented ProCHO5. The transfected cells were incubated over night at 37 °C and 150 rpm in a total volume of 2 ml per well of a 12-well plate. 24 h post transfection the cells were seeded on a 100 mm culture dish in 10 mL CD-Hybridoma medium and cultivated at 37 °C in a humidified atmosphere with 8 % CO_2. To select for positive cell clones 2 mg/mL G418 was added after 5 days. Medium was replaced every 2 - 3 days until the cell clones were picked and transferred into 96-well plates after 2 - 3 weeks. If necessary, the cells were recloned by serial dilution (2.5.8). Clonal cell lines were adapted to suspension cultures in serum free ProCHO5 medium by adding 10 U/mL heparin (Sigma) during the first 2 passages.

2.5.8 Single cell cloning

Subsequent to preparative cell sorting or RMCE clonal cell lines were generated by serial dilutions in a 96-well flat bottom plate or alternatively by seeding 6,000 – 800,000 cells in 60 mm cell culture dishes. For the serial dilution 0.5×10^4 cells/mL in 400 µL ExCell420 medium were inoculated in well A1 and diluted 1:2 along the first column resulting in 200 µL per well. A 1:2 dilution series across all columns was prosecuted also resulting in a total volume of 200 µL per well. In both cases the plates were incubated at 27 °C in a humidified atmosphere until colonies with a diameter of at least 25 cells where clearly visible. Those colonies were drawn into a microtiter pipette tip and passaged into a new 96-well plate. When the cells reached confluency expansion was continued over several passages using 48-, 12-, 6-well plates and 60 mm dishes by detaching with trypsin or a rubber policeman. Finally, cells where transferred into suspension.

2.5.9 Viral titer determination by Plaque Assay

Viral titer were determined using the plaque assay technique. It was originally developed to calculate the titers of bacteriophage stocks and modified by R. Dulbecco in 1952 for use in animal virology (Dulbecco, 1952). A monolayer of susceptible cells is infected by a series of 10-fold dilution of the virus solution. To prevent a distribution of the virus upon cell lysis, the medium is fixed by an overlay of low melting point agarose. Thus, virus particles can only infect neighbouring cells which subsequently leads to the generation of empty spots in the monolayer, the so called plaques. After several days neutral red is added to the monolayer to visualise the plaques caused by viral-dependent cell lysis. The cationic dye binds to negatively charged lysosomal membranes, thus staining only intact cells, whereas plaques are visible as white spots. The virus concentration is calculated in plaque forming units per mL (PFU/mL).

In this work each virus was titered using three dilution stages prepared in ExCell420 ranging from 10^{-4} – 10^{-6} for first generation virus from transfection supernatants and 10^{-5} – 10^{-7} for virus amplificates. In each well of a 6-well plate 1.25×10^6 Sf21 cells were seeded. After the cells have been fully attached the medium was exchanged to remove residual FCS which inhibits virus entry. Subsequently, the cells were covered with 500 µL of the virus dilutions and incubated at 27 °C to allow infection. A virus with known titer and uninfected cells were used as controls. After 2 h the virus solution was exchanged by an overlay of ExCell420 supplemented with 1 % low melting point agarose, 5 % FCS and antibiotics. After the overlay has hardened the plates were incubated for 6 days at 27 °C. Afterwards, each well was stained with 350 µL of a 5 % neutral red solution for at least 4 h. Only those wells with a countable number of plaques between 20 - 200 were used to calculate the virus titer by the equation 3.

$$\frac{\mathbf{PFU}}{\mathbf{mL}} = \frac{\mathbf{average\ no.of\ plaques}}{\mathbf{dilution} \times \mathbf{volume\ of\ virus\ [mL]}} \qquad \text{Eq. 3}$$

2.5.10 Virus amplification

To amplify first generation virus harvested from transfection supernatants, Sf21 cells were seeded in a density of 0.5×10^6 cells/mL in 50 -200 mL serum free ExCell420 and incubated at 27 °C and 100 rpm. The cells were infected using 1 % (v/v) transfection supernatant or with an MOI of 0.1 if the titer was known. Cell number, viability and diameter were determined every 24 h. In case of viruses harbouring fluorescent proteins, the amount of fluorescent cells was measured at the Guava flow cytometer. The supernatant was harvested 48 h after the cells stopped proliferating and reached an average diameter of 20 -25 µm. Following centrifugation for 15 min at 5000 rpm to separate the cells the cleared supernatant was sterilised by filtration using a Stericup vacuum system (Millipore) with a 0.45 µm filter and stored at 4 °C until titer determination.

2.5.11 Cryopreservation of baculoviruses in infected cells

To prevent a loss of infectivity during long term storage, the viruses were preserved using the method of baculovirus infected insect cells (BIIC) (Wasilko *et al.*, 2009). 300 mL of a Sf21 culture with a density of 1×10^6 cells/mL were infected with an MOI of 3 and cultivated at 27 °C and 100 rpm. When a significant increase in cell diameter was observed 24 hpi, the cells were harvested by centrifugation for 4 min at 180 ×g. The infected cells were subsequently resuspended in ExCell420 supplemented with 10 % DMSO and 45 % conditioned medium in a concentration of 1×10^7 cells/mL. Aliquotation in cryo tubes and deep freezing was performed as described in 2.5.4 on page 40.

2.6 Protein production and purification

2.6.1 Plasmid based transient protein expression

Transient expressions of the model proteins were performed to proof the functionality of the vectors pFlpBtM-I and pFlpBtM-II and to evaluate the expression rates of the model proteins in insect cells and HEK293-6E. The transfections were carried out according to 2.5.5 on page 40. Cell number, viability and if applicable the fluorescence were probed every 24 h and samples of the culture supernatant as well as aliquots of 10^6 cells were taken for analysis and stored at -20 °C until further use.

HEK293-6E cultures were expanded 48 h post transfection (hpt) with fresh medium to a total volume of 50 mL and supplemented with a final concentration 0.5 % tryptone (TN1). Additionally, 4.5 g/L glucose were added 72 hpt to compensate the metabolic consumption. To further increase the expression rate after the cell density has reached the stationary phase 96 hpt, 3.75 mM valproic acid (VPA) were added. VPA is a competitive inhibitor of histone deacetylases (HDAC). While it increases the expression rate it has also a negative influence on the proliferation rate, hence it is supplemented only in the late phase of the expression (Backliwal *et al.*, 2008b). Transient expressions in insect cells have not been supplemented and were harvested after 4-7 days as soon as the cell number and viability began to decrease due to nutrient limitations.

2.6.2 Protein expression in insect cell lines using BEVS

For baculoviral protein production insect cells were seeded in 30 - 200 mL suspension cultures with initial cell densities between 0.5×10^6 and 1×10^6 cells/mL. MOIs between 1 - 5 were chosen if titers have been determined before. Otherwise first generation virus stock was added in a total concentration of 10 %. Infection kinetics were monitored every 24 h by the determination of the growth curves, cell diameter and percentage of fluorescent cells. As in plasmid based transient expression, samples of the culture supernatant as well as aliquots of 10^6 cells were taken for expression analysis and stored at -20 °C until further use. The supernatants and pellets were harvested as soon as the viability dropped below 80 % or after 7 days post infection at the latest.

2.6.3 Protein production in stable mammalian cell lines

Small scale protein production in stable CHO Lec3.2.8.1 producer cell lines was performed in shake flasks at 37 °C and 100 rpm in humidified atmosphere with 5 % CO_2. To obtain significant amounts of ECDmTLR2, a cultivation in an autoclavable stirred tank bioreactor with 2.5 L culture volume was performed by Nadine Konisch. The bioreactors were equipped with a double membrane stirrer for both bubble-free aeration via 8 m of hydrophobic polypropylene Accurel™ S6/2 membrane tubing (Membrana) and perfusion with internal cell retention via 8 m of hydrophilised, microporous membrane tubing of the same type (Lehmann *et al.*, 1987; Blasey & Jäger, 1991). The process diagram is shown in Figure 2-2.

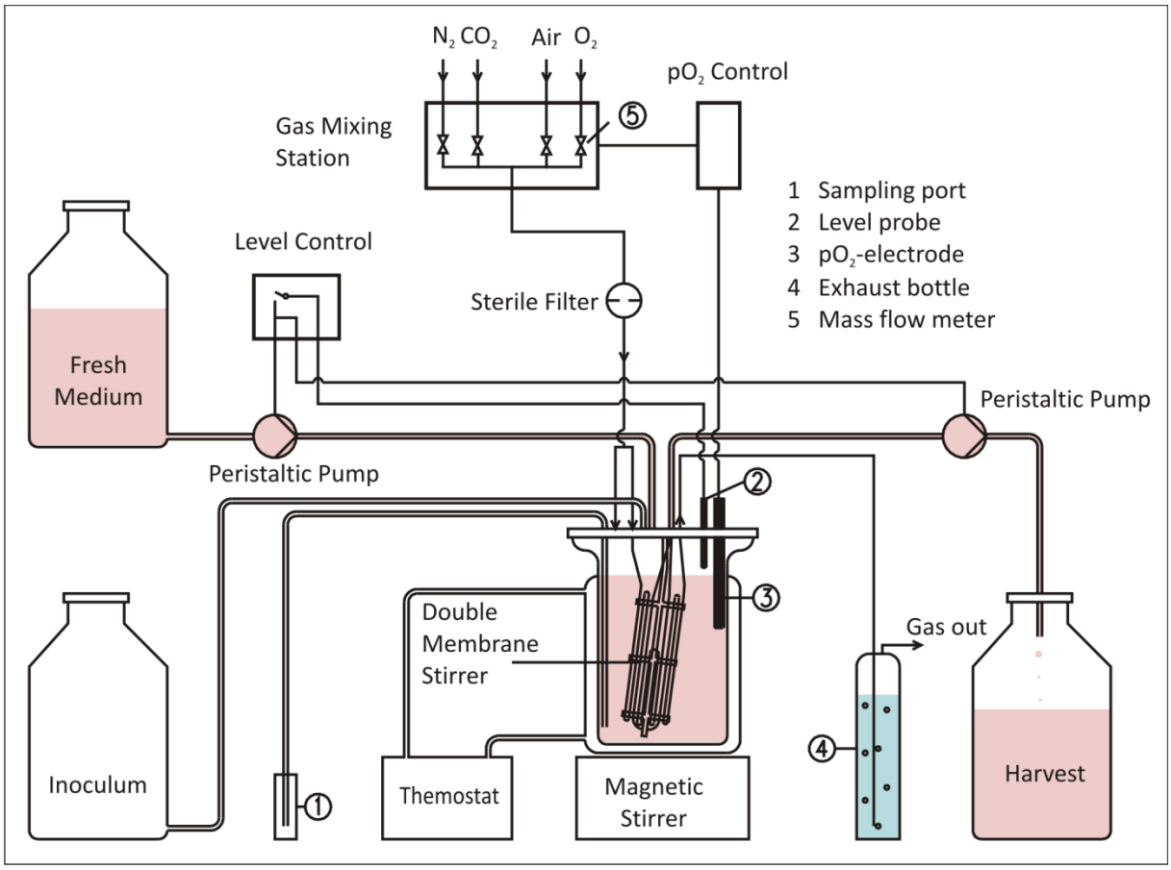


Figure 2-2: Flow diagram of a stirred tank reactor for continuous cultivation of CHO Lec3.2.8.1 producer cell lines
Flow diagram of a continuously perfused double membrane stirred tank reactor with bubble-free aeration for the stable production of recombinant protein by CHO Lec3.2.8.1 producer cells (source: V. Jäger, HZI).

Upon seeding in a 1:1 blend of ProCHO5 and ZKT-I media with an initial cell density of 3×10^5 c/mL the production cell line was propagated in batch mode at 37 °C with a stirring speed of 45 rpm and a dissolved oxygen (DO) concentration of 40 % air saturation at pH 7.2. Upon reaching a cell density of approximately 4×10^6 after 4 days the perfusion mode was started with an initial rate of 2 L/d. During the following cultivation period over 8 days the perfusion rate was gradually increased to a final value of 7.5 L/d according to the metabolic consumption of glucose preventing a drop below 2.5 g/L. At a cell density exceeding 10^7 cells/mL, temperature was reduced to 32 °C. Samples of 2 mL were taken daily for routine in-process control. A total volume of 35 L cell-free supernatant could be obtained from a single production run. Following a 25-fold concentration by ultrafiltration it was diafiltrated against PBS using a Pellicon 2™ tangential flow system (Millipore) equipped with two 30-kDa cut-off cartridges.

2.6.4 Cell lysis

To analyse the expression kinetics aliquots of 1×10^6 cells were lysed by incubation in 150 µL Cytobuster protein extraction reagent (Novagen) for 45 min at RT. To separate soluble and insoluble fractions the lysate was centrifuged for 15 min at 13,000 ×g. Both fractions as well as the culture supernatant were stored at -20 °C.

For large scale extraction of intracellular proteins the cell pellets were resuspended in 25 mL Lysis buffer (Table 2-10) and incubated on ice for 15 -30 min. Optionally, resuspended cells were disrupted by sonication. This was performed in a Sonopuls sonifier (Bandelin) equipped with a TT 13 FZ sonotrode using 80 W in intervals of 1 sec for 2 min. The lysate was centrifugated for 30 min at 30,000 ×g and 4 °C. The clear cell extract was transferred to a new falcon tube and stored at 4 °C until affinity purification.

Table 2-10: Lysis buffer

Composition	Final concentration
Sodium phosphate buffer pH 8	50 mM
NaCl	300 mM
Imidazole	5 mM
IGEPAL CA-630	0,5 %
β-mercaptoethanol	3 mM
DNAse	0,2 µg/mL
Leupeptine (only for lysis of insect cells)	100 µM
Roche complete mini protease inhibitor tabled (EDTA free)	

2.6.5 Affinity chromatography

All model proteins expressed in this work were either equipped with poly histidine tags for capturing by IMAC or expressed as hFC fusions to be purified by Protein A affinity chromatography. For analytical small scale purification of His-tagged proteins Ni-NTA spin columns (QIAGEN) were used according to the manufacturer's protocols. The preparative purification of proteins from culture supernatants or lysates of harvested cells was performed using a Profinia protein purification system (Biorad) with preset run parameters. Columns and buffer compositions are summarised in Table 2-11 and Table 2-12. Prior to loading on the columns supernatants and cell lysates were filtrated using Minisart 0.45 µm syringe filters (Sartorius). To prevent stripping of Ni^{2+} ions from the IMAC column by chelating ions in the insect cell culture medium samples have been dialyzed before loading to the columns.

Table 2-11: Affinity chromatography columns

Column	Column volume	Cat. no.
Bio-Scale Mini Profinity IMAC Cartridge	1 mL	#732-4610
Bio-Scale Mini Affi-Prep Protein A Cartridge	1 mL	#732-4600
Bio-Scale Mini Bio-Gel P6 Desalting Cartridge	10 mL	#732-5304

Table 2-12: Buffers for affinity chromatography

Buffer	Composition
2× Cleaning solution 1	1 M NaCl, 100 mM Tris, pH 8
4× Cleaning solution 2	2 M NaCl, 400 mM NaOAc, pH 4.5
2× Storage solution (desalting column)	4 % benzyl alcohol
2× Storage solution (ProtA column)	20 % ethanol
5× Desalting buffer	Phosphate buffered saline (PBS), pH 7
5× Prot A binding/wash buffer	Phosphate buffered saline (PBS), pH 7
5× Prot A Elution buffer (citrate)	500 mM sodium citrate, pH 3
2× IMAC Wash buffer 1	600 mM KCl, 100 mM KH_2PO_4, 10 mM Imidazole, pH 8
2× IMAC Wash buffer 2	600 mM KCl, 100 mM KH_2PO_4, 20 mM Imidazole, pH 8
2× IMAC Elution buffer	600 mM KCl, 100 mM KH_2PO_4, 500 mM Imidazole, pH 8
Dialysis buffer	50 mM Tris-HCl, 500 mM NaCl, 1 mM $NiCl_2$, 5 mM $CaCl_2$

2.7 Protein analytical methods

2.7.1 Photometric quantification of protein concentrations

Concentrations of purified protein solutions were obtained via the inline measurement of the Profinia device. Alternatively, they were determined using a NanoDrop ND-1000 photometer. The measurement was performed with a sample volume of 2 µL after a calibration with a corresponding protein free buffer.

Both the Profinia and the NanoDrop device indicate the absorbance scaled to a layer thickness of 1 cm. According to the Beer-Lambert law (Eq. 4) the absorbance value had to be divided by the mass extinction coefficient to calculate proper protein concentrations. Since the extinction at a wavelength of 280 nm is caused by the aromatic amino acids tryptophan and tyrosine, protein specific mass extinction coefficients were calculated for each protein by the VectorNTI software (Invitrogen).

$$c \ \frac{mg}{mL} = \frac{A_{280}}{\varepsilon \times l} \qquad \text{Eq. 4}$$

c: protein concentration
A_{280} dimensionless absorbance at 280 nm
ε mass extinction coefficient [mL · mg^{-1} · cm^{-1}]
l: layer thickness [cm]

2.7.2 Fluorometric quantification of eGFP concentrations

A quantification of intracellular eGFP accumulation was performed by measuring fluorescence of cell lysates at a NanoQuant InfiniteM200 plate reader (Tecan). For an accurate acquisition a serial dilution of the cell lysates from 1:2 to 1:256 was prepared. An eGFP standard (Biovision) was used to set up a calibration curve ranging from 0.1 µg/L to 1 µg/L to allow a quantification of protein yields from relative fluorescence intensities. The measurements were performed as duplicates in black F96 MicroWell™ plates (Nunc) using the parameters summarised in Table 2-13.

Table 2-13: Plate reader settings

Parameter	Value
Shaking duration	2 s
Shaking amplitude	2 mm
Mode	Fluorescence top reading
Excitation wavelength	488 nm
Emission wavelength	509 nm
Excitation bandwidth	5 nm
Emission bandwidth	5 nm
Gain	10 Manual
Number of flashes	50
Flash frequency	400 Hz
Integration time	20 µs
Lag time	0 µs
Settle time	0 ms
Z-Position (manual)	20000 µm

2.7.3 SDS-PAGE

Denaturing polyacrylamide gel electrophoresis (Laemmli, 1970) was used to determine expression kinetics during the protein production as well as to evaluate purity and yield of purified proteins. Either Criterion™ precast gels (Biorad) or self-made gels composed of a lower resolving gel (12 % acrylamide) and a upper stacking gel (5 % acrylamide) were used (Table 2-14). Typically, 10 - 15 µL of the protein sample were mixed with 2 µL of a 8× SDS loading buffer and heated for 5 min at 90 °C. The electrophoresis was performed at a constant voltage of 160 V until the dye front reached the end of the gel. Subsequently, the gel was washed with water to remove SDS and then stained with Instant Blue (Expedion) or a methanol based coomassie dye. The buffers used for SDS-PAGE are summarised in Table 2-15.

Table 2-14: Compositions of SDS-PAGE gels (Volumes for 4 gels)

Solution	12 % resolving gel	5 % stacking gel
Acrylamide/bisacrylamide 30 % (v/v)	8.0 mL	0,75 mL
4× lower buffer	5.0 mL	-
4× upper buffer	-	1.25 mL
10 % SDS	0.2 mL	-
H_2O	6.8 mL	2.95 mL
TEMED	20 µL	15 µL
40 % APS	30 µL	15 µL

Table 2-15: SDS-PAGE buffers and reagents

Buffer	Composition
4× Upper buffer	0.5 M Tris-HCl pH 6.8, 0.4 % SDS (v/v)
4× Lower buffer	1.08 M Tris-base, 0.42 M Tris-HCl, pH 8.8
SDS running buffer	3 g/L Tris-base, 14.4 g/L Glycin, 1 g/L SDS, pH 8.3
Coomassie staining solution	50 % Ethanol, 10 % acetic acid, 0.25 % Coomassie R-250
Coomassie destaining solution	50 % Ethanol, 10 % acetic acid

2.7.4 Native PAGE

In a native polyacrylamide gel electrophoresis protein samples are separated without denaturing the secondary, tertiary and quaternary structures. Thus, the migration speed is not only dependent on the protein mass, but also influenced by net charge of the protein surface (determined by the isoelectric point and the pH of the buffer) and the hydrodynamic radius (i.e. the shape and conformation of the protein). To ensure that the proteins migrate to the cathode, a slightly negative charge is mediated by Coomassie G-250 without denaturing the proteins.

In this work the NativePAGE Novex 4-16 % Bis-Tris gel system kit (Invitrogen) was used which is based on the principle of Blue Native PAGE (Schagger & von Jagow, 1991). Since the migration speed is influenced by ion strength of the buffer (Gallagher, 2001), a buffer exchange against 50 mM Tris was performed using Vivaspin sample concentrators (GE Healthcare). As a size marker, the HMW Native (66-669 kDa) standard of the HMW Calibration Kit for native PAGE (Amersham) was used. The electrophoresis was performed according to the manufacturer's instructions at a constant voltage of 150 V for 120 min.

2.7.5 Western blot

For a specific detection of proteins Western blots followed by immunostaining were performed. A BioRad Trans-Blot-SD device was used to transfer protein samples separated in SDS-PAGE to a polyvinylidene difluoride membrane (Immobilon-P, Millipore) by a semi dry blotting procedure (Kyhse-Andersen, 1984). The buffers and antibodies used for this method are summarised in Table 2-16. The gels and 2 pieces of gel-sized Whatman paper were equilibrated for 10 min in transfer buffer. The PVDF-membrane was activated in 100 % methanol for 5 sec, rinsed with water and briefly equilibrated in transfer buffer for 10 min. Subsequently, the gel was placed onto the

membrane avoiding air bubbles to be enclosed. For the transfer, both were placed between two layers of soaked Whatman paper onto the anode of the blot apparatus. The blot was run with 14 V for 30 min. To saturate free binding sites the membrane was incubated for 1.5 h in a blocking solution of 5 – 10 % skim milk powder in TBS-T. Following a washing procedure in TBS-T (3 × 5 min), the membrane was incubated with antibodies against polyhistidine- or hFc-tags overnight at 4 °C. Prior to the incubation in with the secondary antibody, another washing step was performed. If needed, 1 % skim milk powder was added to the antibody solutions to prevent unspecific binding. To visualise the proteins specifically bound by alkaline phosphatase conjugated antibodies the BCIP/NBT Color Development Substrate from Promega was used according to the manufacturer's guidelines.

Table 2-16: Buffers and reagents for Western blots

Buffer	Composition
Transfer buffer	25 mM Tris-base, 192 mM Glycin, 15 % Methanol, pH 8
TBS-T	20 mM Tris-base, 150 mM NaCl, 0.05 % (v/v) Tween, pH 8
AP-buffer	100 mM Tris 100 mM NaCl, 5 mM $MgCl_2$, pH 9.5
BCIP	50 mg/mL BCIP in 100 % dimethylformamide
NBT	50 mg/mL NBT in 70 % dimethylformamide.

Antibody	Type	Concentration	Supplier
His-Tag antibody (mouse)	primary	1:2000 in TBS-T	Novagen #70796
Goat-anti-human IgG (H+L)-AP	primary	1:5000 in TBS-T	Promega #S3821
*Strep*MAB-Classic (mouse)	primary	200 ng/mL in PBS + 0,2% BSA	IBA 2-1507-001
Goat-anti-mouse IgG (H+L)-AP	secondary	1:7500 in TBS-T	Promega #S372B

2.7.6 MALDI-TOF

MALDI-TOF (Matrix Assisted Laser Desorption Ionisation - Time-Of-Flight) mass spectrometry (MS) was used to identify proteins fragments of SDS-PAGE samples. After a tryptic digest, proteins were co-crystallised with organic acids. Desorption of protein fragments is triggered by a UV laser beam followed by an ionization mediated by protonation of the peptides. Time-of-flight measurements reveal mass-to-charge ratio of the ions. This information can be used for protein identification by comparing the data with the MASCOT database. The analyses were carried out in the research group Cellular Proteome Research at the HZI.

2.7.7 N-terminal sequencing

For a detailed characterization of protein fragments the N-termini were sequenced by automated Edman degradation (Edman & Begg, 1967). The protein samples were separated by SDS-PAGE and transferred onto a PVDF membrane specialised for sequencing purposes (Immobilon PSQ, Millipore) as described in 2.7.5 on page 54. Protein bands were cut out after visualization by Coomassie staining. The N-terminal sequencing was performed using the 494A HT Protein Sequencer (Applied Biosystems) by Beate Jaschok-Kentner in the department of Molecular Structural Biology at the HZI.

3 Results

The initial design of a multi-purpose vector aimed at its use for transient expression in insect cells as well as a donor vector for the BEVS and for RMCE in CHO Lec3.2.8.1 master cell lines. Thus, it was named pFlp-Bac-to-Mam (here referred to as pFlpBtM-I) to emphasise its properties as a shuttle vector for baculoviral and mammalian expression by Flp mediated cassette exchange. The construction and evaluation of this vector is presented in chapters 3.1 and 3.2. For the proof of concept studies, protein expression was determined by analytical SDS-PAGE, Western blots and by determining fluorescence via flow cytometry and fluorescent microscopy. A second improved version (pFlpBtM-II) was generated by the incorporation of genetic elements for transient expression in HEK293-6E cells (chapter 3.3). Subsequent to the proof of concept (chapter 3.4) the vector was used for comparative expression and purification studies of model proteins in different hosts (chapter 3.5). Finally, the integration of a RMCE based system for stable expression of heterologous proteins in Sf21 and Hi5 cells is presented in chapter 3.6.

3.1 Construction of pFlpBtM-I

The generation of pFlpBtM-I followed a three-step modification of an in-house variant of the pFastBac donor vector (Invitrogen) as shown in Figure 3-1. Firstly, the polyhedrin promoter of pFastbac-H8PPFlag was replaced by a 1294 bp PCR-fragment of pIEx/Bac5 (Novagen) upon a BbsI-NsiI digest. It comprises the *Ac*MNPV immediate early promoter ie1 and its enhancer hr5 for transient expression and the very late baculoviral p10 promoter. The resulting plasmid pFB-ie10-H8PPFlag was confirmed by analytical restriction digest and sequencing of the corresponding plasmid segment. Subsequently, the BbsI site was eliminated by site directed mutagenesis, as this enzyme should be used for the seamless integration of target gene sequences into to MCS of the final vector. The elimination of the restriction site was again checked by analytical restriction digest and sequencing. Finally, a 1285 bp PCR-fragment containing the complete FRT-cassette from pFS-sighis-PGKd flanked by 5'-BamHI and 3'-AvrII was generated. The parental pFS-sighis-PGKd is a mutant of GenBank JF313343, in which the BamHI site is eliminated via fill in and ligate reactions. The PCR-product was integrated into pFB-ie10-H8PPFlag upon digestion with BamHI-AvrII, thereby eliminating the original MCS in the target plasmid.

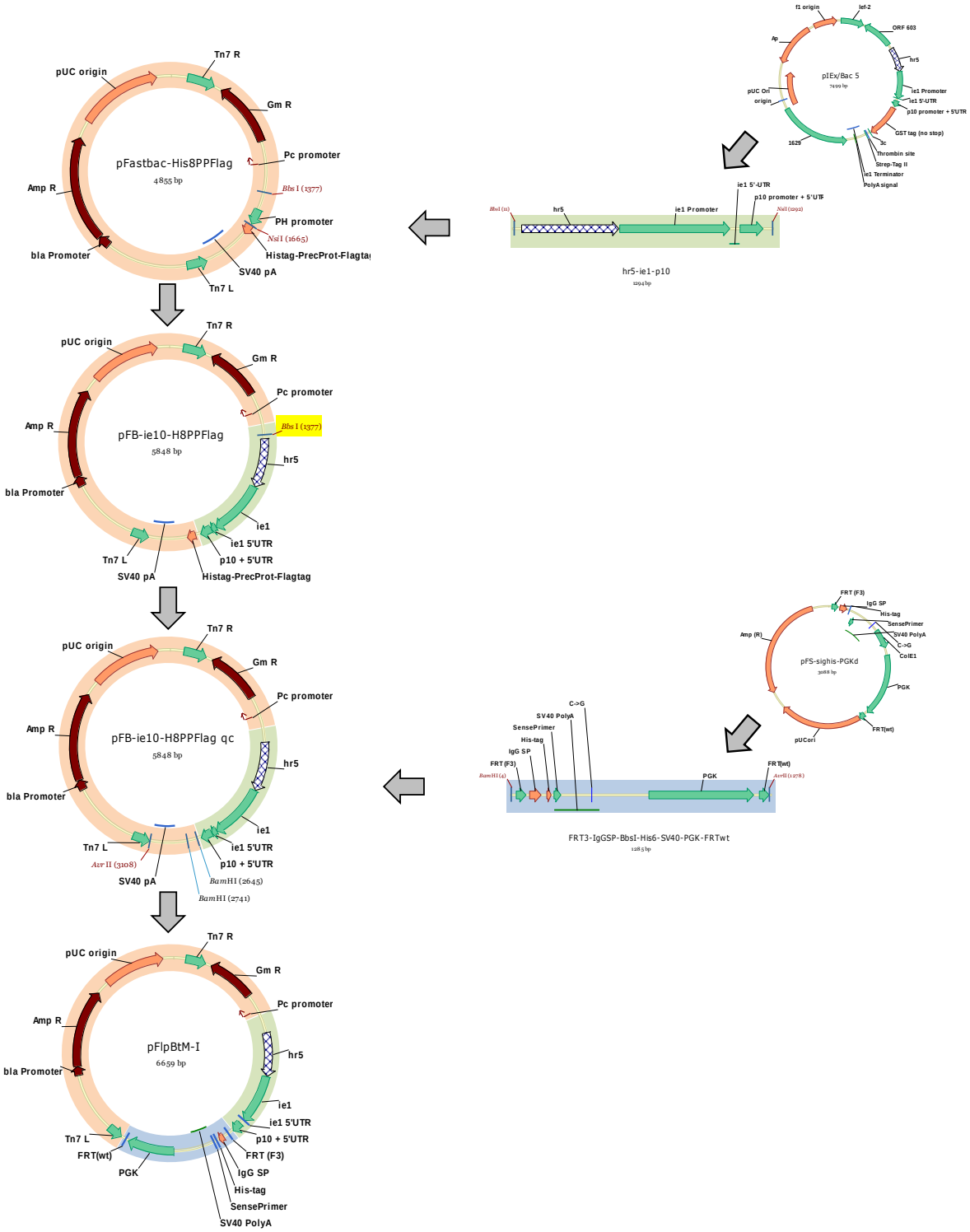


Figure 3-1: Cloning strategy for the construction of pFlpBtM-I
A fragment containing the baculoviral immediate early promoter 1 (ie1) and the homology region 5 enhancer (Hr5) as well as the very late p10 promoter was amplified by PCR from pIEx/Bac5 and integrated into pFastBac-H8PPFlag upon deletion of the original polyhedrin promoter by BbsI-NsiI digest. Subsequently, the BbsI site in the resulting intermediate pFB-ie10-His8PPFlag was eliminated by a site directed mutagenesis. The MCS of the pFB-ie10-His8PPFlag_qc mutant (qc = Quickchange) ranging from BamHI-AvrII was exchanged by the RMCE cassette comprising a pair of heterologous FRT sites, the PGK promoter and a start codon for the selection trap and a MCS with an IgG signal peptide, a C-terminal His_6 tag and the SV40 polyA. The cassette was generated as a PCR-fragment with flanking restriction sites.

The RMCE cassette contains an upstream FRT3-mutant, an IgG signal peptide, two opponent BbsI-sites, a C-terminal His_6 tag and a SV40 polyadenylation signal (pA). Additionally, the PGK promoter and the ATG start codon for the selection trap as well as the second FRTwt site are located within the cassette downstream of the SV40 pA element. The resulting pFlpBtM-I was again confirmed by sequence analysis and analytical restriction digest. Figure 3-2 shows a schematic map of the final construct composed of the three sections from different parental plasmids.

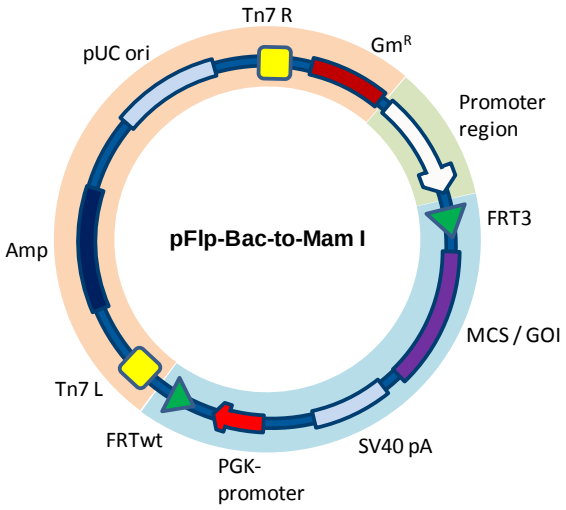


Figure 3-2: Schematic map of pFlp-Bac-to-Mam-I
Vector map of pFlpBtM-I with elements for recombination, expression, screening and replication. Sections originating from different parental vectors are highlighted in orange (pFastBac), blue (pFS-sighis-PGKd) and green (pIEx/Bac5), respectively. The promoter region contains the baculoviral immediate early ie1 and late p10 promoters.

The design of the vector is based on the overlap of the regions for Tn7 transposition and RMCE. The Tn7 sites which mediate the transposition into the baculoviral genome, flank a 3576 bp large region containing a Gentamicin resistance for the screening of positive integrations into the bacmid DNA, the promoter region, the MCS and additional elements for RMCE. Those are the PGK promoter, the ATG start codon for the selection trap and the FRT sites upstream of the MCS and directly behind the ATG. Thus, no prokaryotic or baculoviral elements are integrated into the host cell genome upon RMCE, whereas the elements needed for the RMCE will be integrated into the bacmid DNA upon transposition (Figure 3-3). Furthermore, the integration of the baculoviral immediate-early promoter 1 enables the use as an expression vector for direct transient expression in Hi5 or *Spodoptera* derived cell lines.

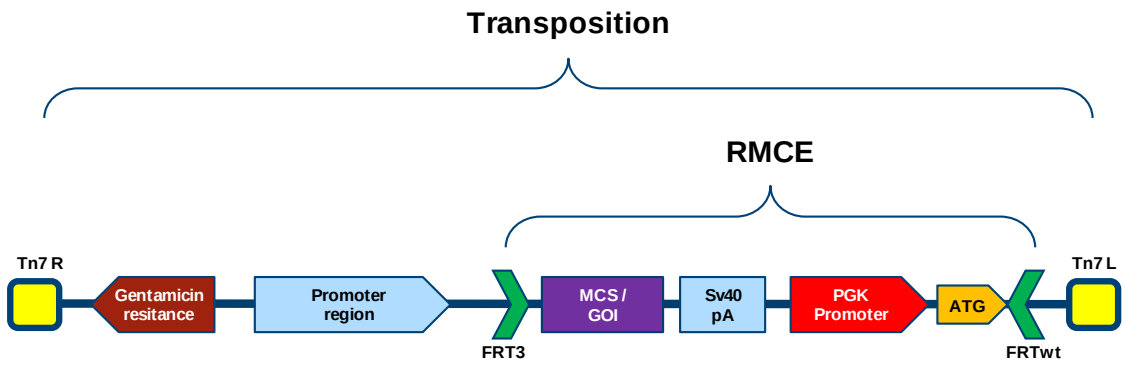


Figure 3-3: Regions for transposition and RMCE in pFlp-Bac-to-Mam
The multi-purpose vector pFlpBtM contains FRT sites for the use as a donor vector for Flp-recombinase mediated cassette exchange and Tn7-transposition sequences for the generation of recombinant bacmids. The PGK promoter is used for a selection-trap method to screen for positive clones. FRT sites flank the region, which is integrated into the RMCE *locus* in the master cell line. It contains the MCS and a downstream PGK promoter for the selection trap method to screen for recombined cell clones. A larger section of the plasmid comprising a Gentamicin-resistance gene for the selection of recombinant bacmids and the promoter region is integrated into the bacmids by Tn7-based transposition. The promoter region contains the early viral ie1 promoter and the hr5 enhancer which drive constitutive expression in Sf21 and Hi5 as well as the very late p10 promoter which is only active in the late phase of viral infection.

3.2 Evaluation of pFlpBtM-I

To determine the applicability of the constructed vector for transient, viral and stable expression, test expressions of fluorescent model proteins have been performed. Therefore, genes of eGFP (gb ADD98904.1), mCherry (gb AY678264) and the ECD of mTLR2 were integrated into the plasmid. A set of opposing non palindromic Type IIS endonucleases for the integration of target proteins facilitate seamless in frame fusion of the target gene to the N-terminal IgG signal peptide and the C-terminal polyhistidine tag of the vector. By attaching BbsI sites, or any other type IIS restriction site, to the gene of interest by a PCR amplification any desired overhang can be generated for the seamless integration. However, for the intracellular accumulation of target proteins the corresponding gene has to be integrated into pFlpBtM through a PCR-fragment flanked by a 5'-NcoI site. By cutting the pFlpBtM vector with NcoI the IgG signal peptide is excised. The same strategy has been used for ECDmTLR2 which was integrated with its authentic signal peptide. A detailed presentation of the MCS and the integration strategy is shown in Figure 3-4. Proper integration of the model proteins has been confirmed by analytical restriction digests and sequence analyses.

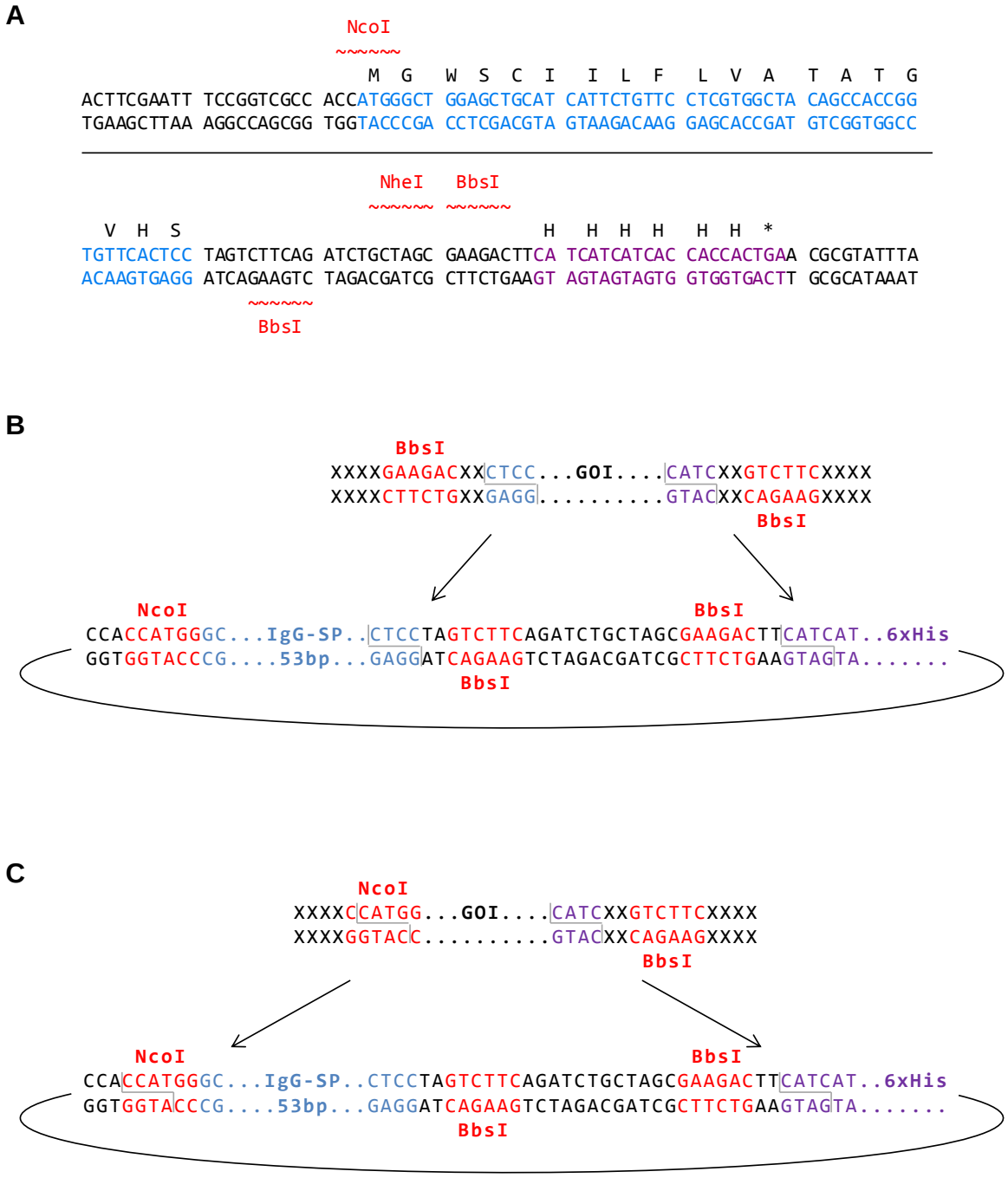


Figure 3-4: MCS of pFlp-Bac-to-Mam-I
A: Sequence of the MCS of pFlpBtM-I. The IgG signal peptide sequence is indicated in blue, the translation is stated above. The two opposing BbsI recognition sites are marked in red and their resulting overhangs are indicated by the grey lines. **B:** Seamless integration of genes can be performed using the two BbsI sites since they are non-palindromic Type IIS endonucleases. This integration strategy generates an in-frame fusion to the signal peptide and the His_6 tag. For this method PCR-primers attaching any Type IIS restriction site of choice are necessary that append compatible overhangs to the gene of interest. The resulting overhangs are part of the signal peptide (blue) and the His_6 (purple), respectively. **C:** A 5'- NcoI site can be used for the deletion of the signal peptide. The 3' integration is the same as described above. GOI = gene of interest, IgG-SP = signal peptide, n,x = arbitrary bases, black lines indicate the cutting position and resulting overhangs of restriction nucleases.

3.2.1 Transient expression in Sf21

Plasmid based transient transfection of insect cell cultures has not been established in the PSPF before. Therefore, initial screenings to optimise transfection efficiencies by comparative test expressions with different transfection reagents, conditions and cell lines had to be performed. Moreover, preceding experiments had to be carried out to facilitate an efficient downscaling to prosecute transfection in small scale cultures of less than 15 mL for economic reasons. The optimisations of the culture conditions and transfection protocols as well as subsequent test expressions were carried out with technical assistance of Carmen Lorenz during her practical term under the author's direction. Following the results of the optimisations, transfections of Sf21 with plasmid DNA were performed in 4 - 10 mL cultures with an initial cell density of 1×10^6 cells/mL in TPP bioreactor tubes. A DNA dose of 1.5 µg per 1×10^6 cells was used with a DNA/Lipofectin ratio of 1:2. Transfection rates of more than 50 % were confirmed by flow cytometric analyses (data not shown). However, only weak fluorescence could be observed by fluorescence microscopy (Figure 3-5).

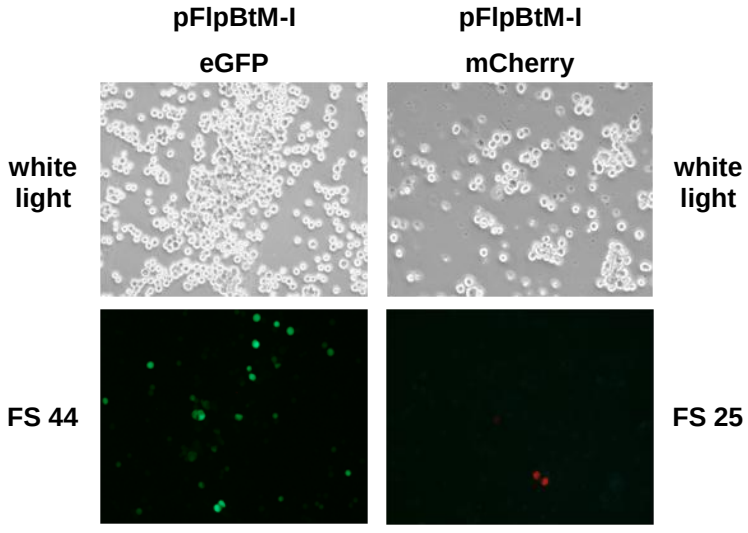


Figure 3-5: Fluorescence microscopy of Sf21 transfected with pFlpBtM-I-mCherry and –eGFP
Pictures were taken 48 h post transfection under white light and illuminated by a HBO50 mercury-vapor lamp using Zeiss filter sets 44 (eGFP) and 25 (mCherry), respectively. The exposure time for fluorescence detection was 6 s.

The expression efficiency was additionally assessed by SDS-PAGE and Western blotting of cell lysates and culture supernatants. Coomassie stained gels and Western blots after detection with an α-His primary antibody of the expression of mCherry and ECDmTLR2 are shown in Figure 3-6.

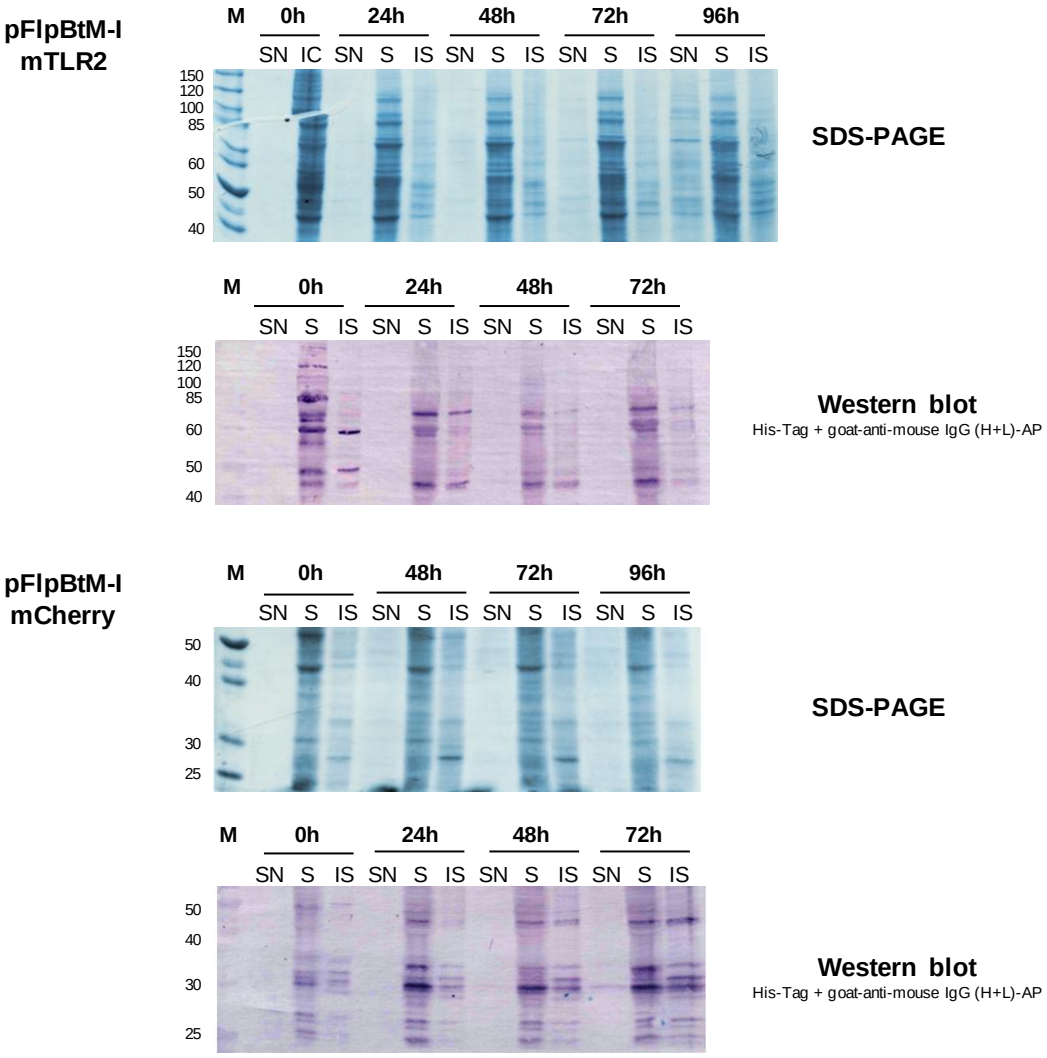


Figure 3-6: SDS-PAGE and Western blots of intracellular fractions and supernatants of transfected Sf21
Samples of non-transfected cells (0h) and between 24 h and 96 h post transfection were taken. The culture supernatant (SN) as well as cell lysates (IC) or soluble (S) and insoluble (IS) intracellular fractions were analysed by SDS-PAGE. Western blots were incubated with a His-Tag primary antibody and a goat-anti-mouse IgG (H+L)-AP conjugated secondary antibody. The model proteins should give signals at 28 kDa (mCherry) and 67 kDa (ECDmTLR2), respectively. M = PageRuler™ Protein Ladder [kDa]

Notably, although expression of the fluorescent model proteins could be detected by flow cytometry and fluorescence microscopy, no bands of the heterologous proteins could be identified on the SDS-PAGE. Moreover, no explicit signals of the target proteins were achieved by immunostaining despite extended incubation with the detection substrates. Signals which might correlate with the expected protein sizes are relatively faint compared to unspecific signals of host cell proteins which are a result of the rather low specificity of the α-His antibody. To determine whether the low expression rate is caused by either the expression method, the vector architecture or a specific characteristic of the model proteins, test expressions in the BEVS were performed.

3.2.2 Generation of recombinant bacmids and baculoviral expression

The identical pFlpBtM-I constructs harbouring eGFP, mCherry and ECDmTLR2 respectively that were used in transient expression have been employed to generate recombinant bacmids. For the expression of mCherry and ECDmTLR2, host strains containing an EMBacY bacmid have been used which harbours an YFP gene for monitoring infection kinetics. In contrary, pFlpBtM-I-eGFP was integrated into conventional MultiBac bacmids since fluorescence of eGFP and the bacmid encoded YFP would not be distinguishable. Blue-white selection of positive clones, the isolation of recombinant bacmid DNA and confirmation of integrations by analytical PCR were conducted according to the MultiBac manual. Upon transfection of adherent Sf21 cells with the generated bacmids the resulting baculoviral expression of the model proteins was confirmed by fluorescence microscopy (Figure 3-7).

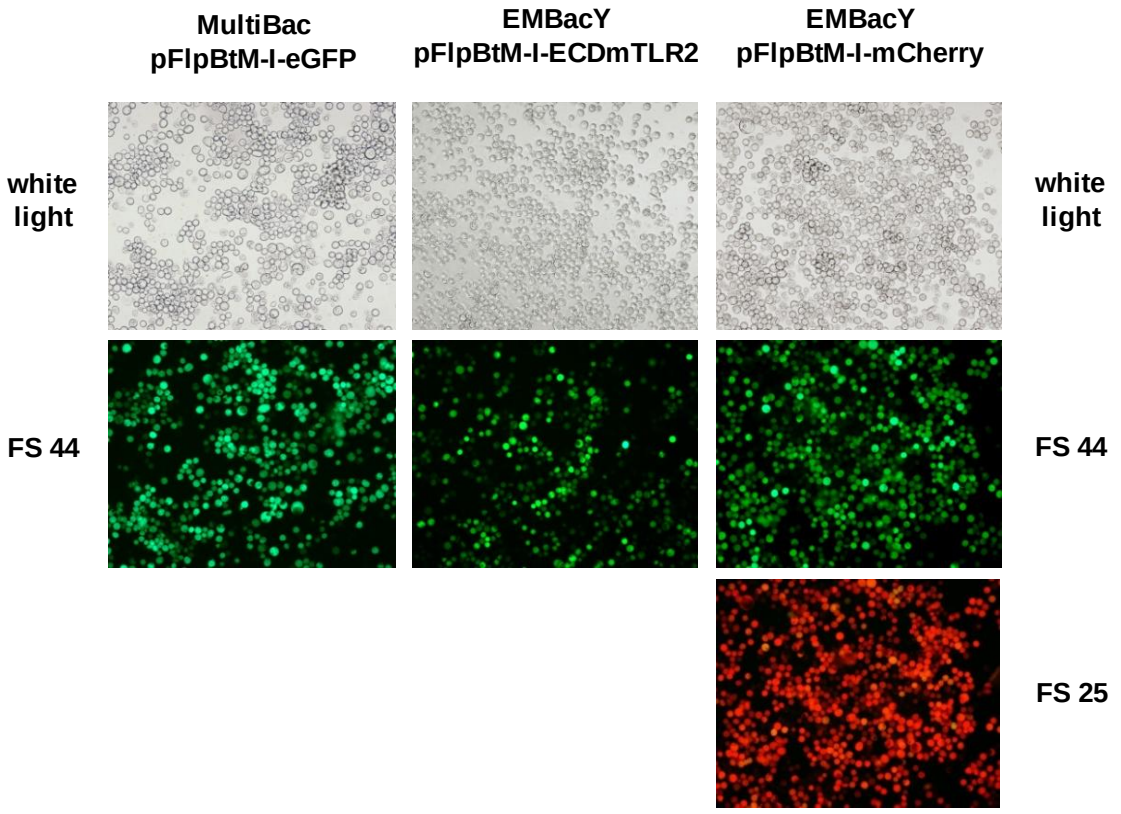


Figure 3-7: Fluorescence microscopy of Sf21 transfected with pFlpBtM-I derived recombinant bacmid DNA
Pictures were taken 9 days post transfection under white light and illuminated by a HBO50 mercury-vapor lamp using Zeiss filter sets 44 (eGFP/YFP) and 25 (mCherry), respectively. The exposure time for fluorescence detection was 1 s. EMBacY variants show fluorescence in the green filter due to the encoded YFP. Cells infected by EMBacY-pFlpBtM-I-mCherry and MultiBac-pFlpBtM-I-eGFP produce significant amounts of the reporter genes. Fluorescence of eGFP is brighter than YFP due to the specifications of the filter set and the higher quantum yield of the protein.

To verify the infectivity and expression capability of fully assembled viruses, large scale test expressions were performed following a virus amplification of the harvested transfection supernatants. A titer determination by plaque assay allowed using a specific MOI of 2. The growth kinetics and variations in cell diameter of the corresponding test expressions are shown in Figure 3-8.

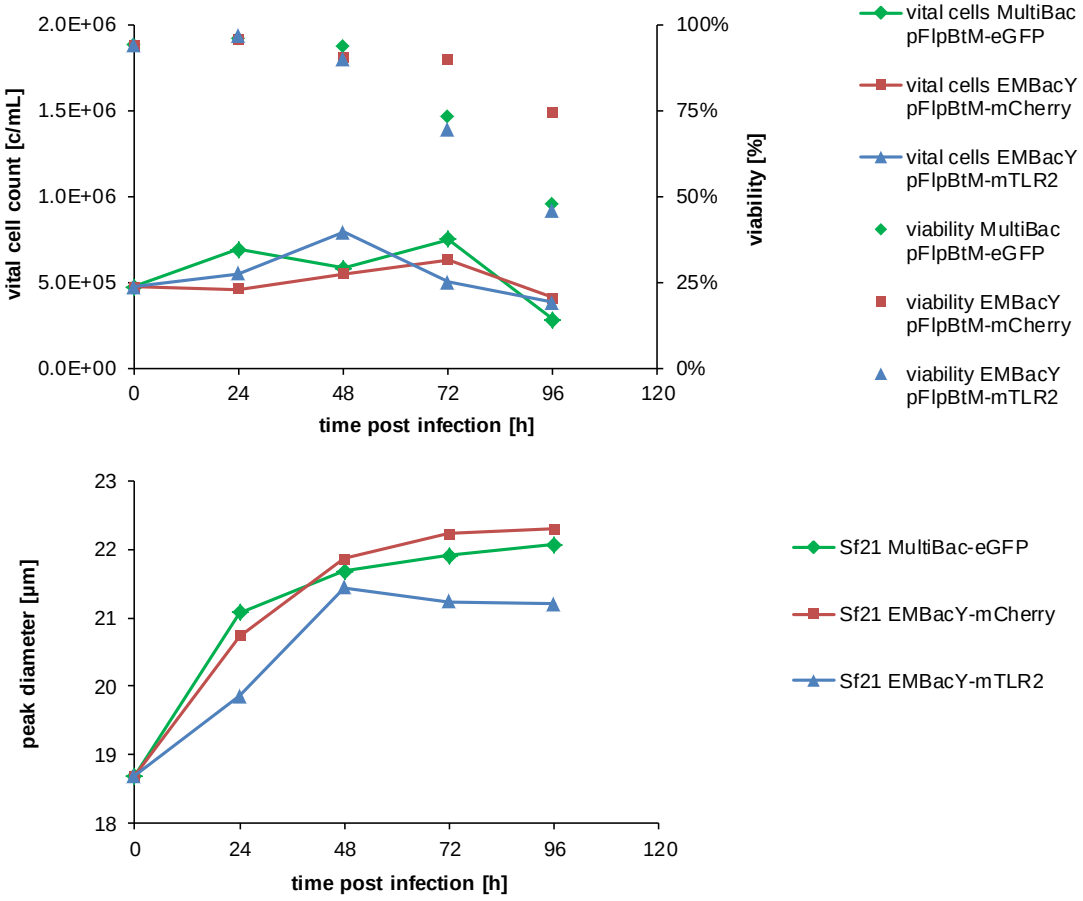


Figure 3-8: Growth kinetics and cell diameter of Sf21 cells infected with pFlpBtM-I derived baculoviruses
0.5×10^5 cells/mL were infected with an MOI of 2. Significant increase in cell diameter of approximately 20 % and initial stop of proliferation upon transfection indicate a successful infection. Consequently, a decline of viability beginning 48 hpi is observed.

The kinetics of both the virus amplification (data not shown) and the expression indicate that the viruses which were generated with pFlpBtM-I as donor vector are fully infective, as proliferation stopped immediately upon infection. The expression and intracellular accumulation of the fluorescent reporter proteins was clearly visible by the staining of the culture. By analogy to the transient expression in Sf21 SDS-PAGE and immunoblotting of

cell lysates and culture supernatants were performed and exemplary shown for mCherry and ECDmTLR2 in Figure 3-9.

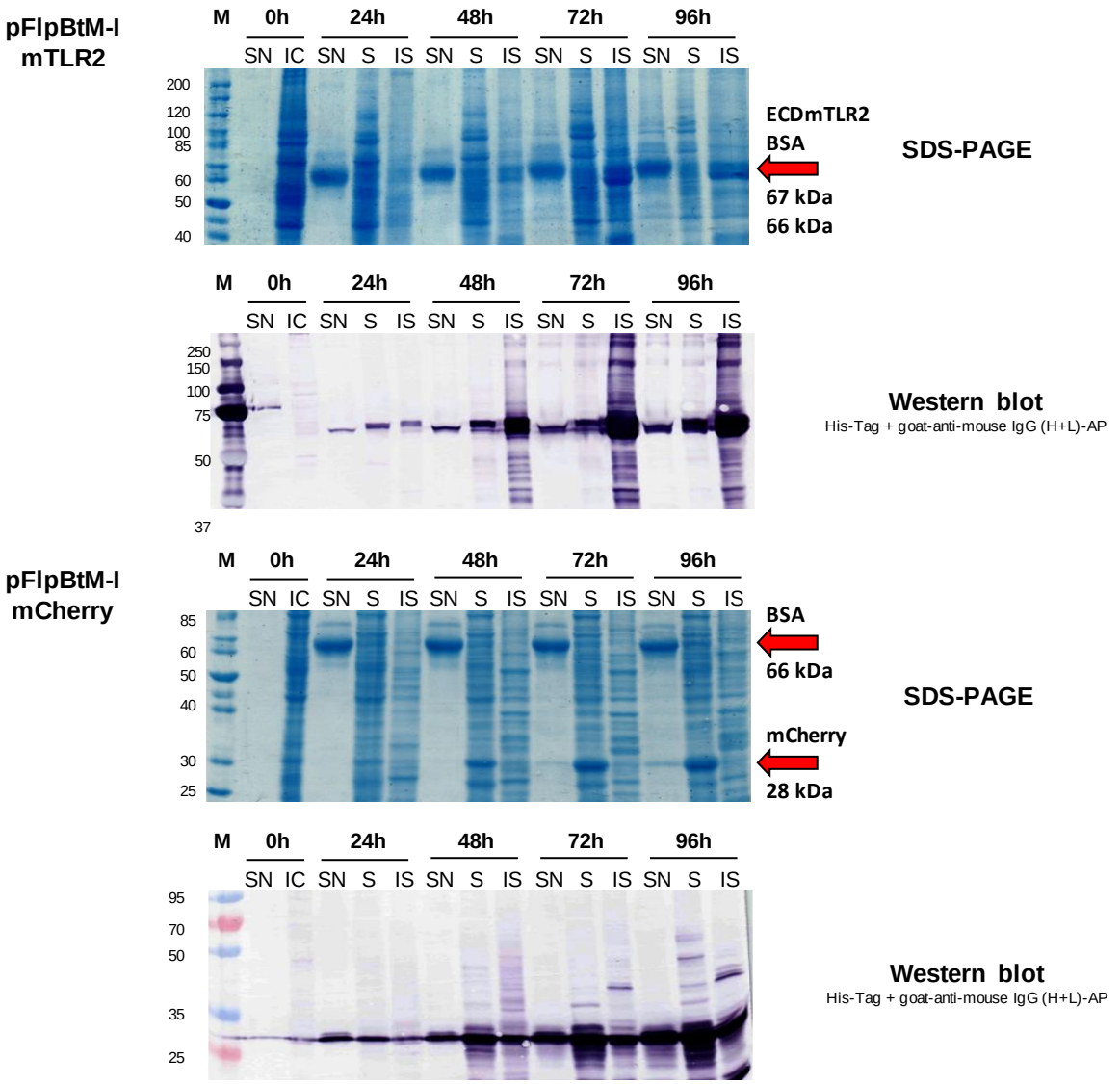


Figure 3-9: SDS-PAGE and Western blots of intracellular fractions and supernatants of Sf21 infected with recombinant baculoviruses generated with pFlpBtM-I as donor vector
Samples of culture supernatant (SN) as well as cell lysates (IC), soluble (S) and insoluble (IS) intracellular fractions from non-transfected cells (0h) and from sampling between 24 h and 96 hpi were analysed. Western blots were incubated with a His-Tag primary antibody and a goat-anti-mouse IgG (H+L)-AP conjugated secondary antibody. The model proteins should give signals at 28 kDa (mCherry) and 67 kDa (ECDmTLR2), respectively. In both expressions strong bands of BSA at 66 kDa are present. M = PageRuler™ or PageRuler™ Plus Prestained Protein Ladder [kDa]

Intracellular accumulation of soluble mCherry could be validated by coomassie stained gels as well as by specific signals on Western blots. Likewise, characteristic signals of eGFP were seen on the corresponding gels and blots (data not shown). In contrary,

expression of ECDmTLR2 could not be clearly identified on SDS-PAGE since serum protein from the virus supernatant is superposing the expected size range of the ECDmTLR2 of 67 kDa (MW of BSA = 66.4 kDa). However, Western blot analysis undisputedly proved the expression of the model protein. Conspicuously, although the protein contains a signal peptide and thus should be secreted, the predominant amount of the recombinant protein remains in the insoluble fraction. Despite the protein specific occurrence of insoluble material the production of all three model proteins was clearly demonstrated using pFlpBtM-I derived recombinant viruses. Thus, the plasmid's suitability to be used as a donor vector for baculoviral expression in insect cells could be confirmed. The next step was consequently to demonstrate the applicability of pFlpBtM-I as a donor vector for RMCE.

3.2.3 Flp-recombinase mediated cassette exchange

Due to the labour intensive clonal isolation procedure, the cassette exchange was tested with only one of the model proteins. The construct pFlpBtM-I-mCherry was chosen since its expression upon stable integration in the genuine CHO Lec3.2.8.1 production cell line can be detected most easily. Upon transfection of the master cell clone SWI3-26 with pFlpBtM-I-mCherry, G418 resistant cell clones were isolated and propagated as clonal isolates. The expression of the reporter protein in the resulting producer cell line was confirmed by fluorescence microscopy as shown in Figure 3-10. Stability and homogeneity of the genomic protein expression was additionally analysed by flow cytometry. Figure 3-11 shows the fluorescence profiles of the parental master cell line SWI3-26 and the final producer clone 12 weeks post transfection.

The data confirmed the microscopic observation of homogeneous expression of mCherry throughout the population. Additionally, no differences in recombination efficiency as well as in the robustness and stability between the producer cell lines generated with conventional targeting vectors or pFlpBtM-I could be observed. Accordingly, the results demonstrate that the new multipurpose vector can be used for RMCE without any constraints.

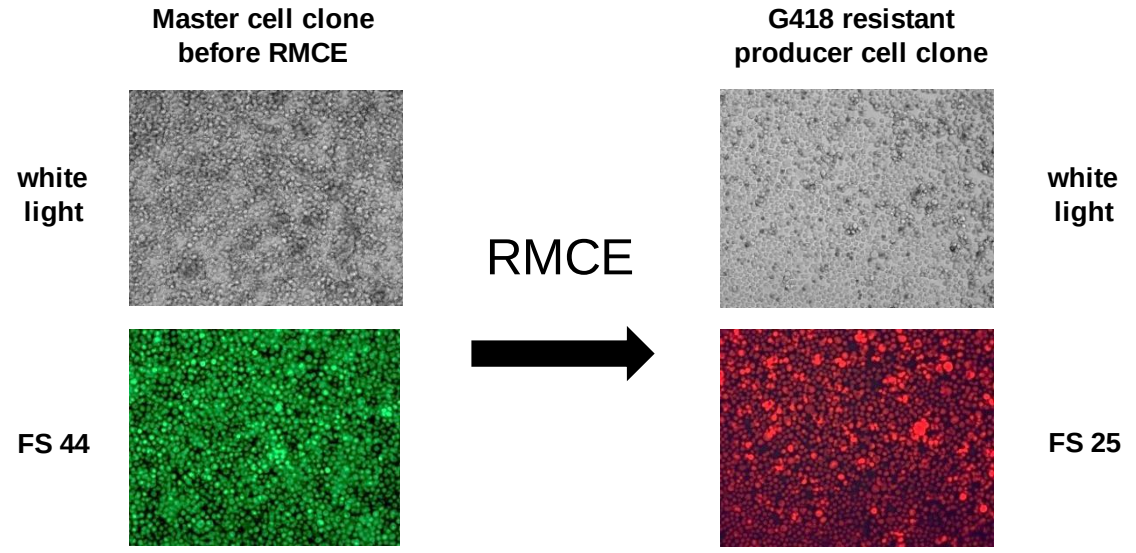


Figure 3-10: CHO Lec3.2.8.1 RMCE master cell line before and after cassette exchange with pFlpBtM-I-mCherry
The RMCE master cell line CHO Lec3.2.8.1 SWI3-26 homogeneously expresses the reporter protein eGFP (left). Upon cassette exchange with pFlpBtM-I-mCherry, the resulting producer cell clone stably expresses mCherry with no eGFP background (right). The exposure time for fluorescence detection was 1 s.

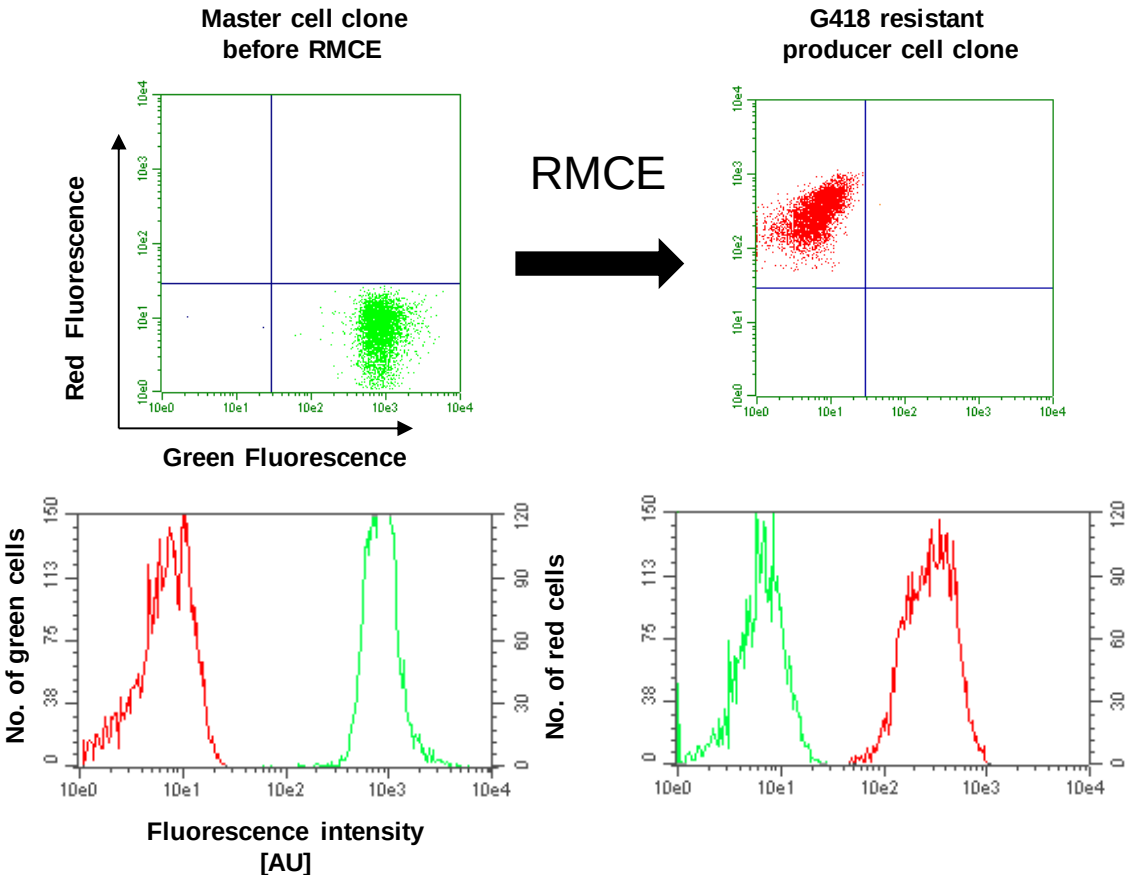


Figure 3-11: Dot plots and histograms of CHO Lec3.2.8.1 before and after RMCE with pFlpBtM-I-mCherry
Fluorescence of the CHO Lec3.2.8.1 RMCE master cell line SWI3-26 (left) and the producer cell line after the cassette exchange with pFlpBtM-I-mCherry (right) was assessed at the Guava EasyCyte Mini cytometer with an excitation wavelength of 488 nm. Both dot plots (upper row) and histograms (lower row) demonstrate the exchange of eGFP by the reporter mCherry.

3.3 Construction of pFlpBtM-II

The results of the expression studies with pFlpBtM-I demonstrated that plasmid based transient expression in Sf21 is possible, but protein yields are insufficient for preparative protein production for crystallization or biochemical analyses. Yet, the plasmid proved to be particularly suitable to act as a donor for RMCE and the BEVS. The construction of an improved second generation of pFlpBtM was therefore performed by the integration of genetic elements conferring applicability for transient expression in the efficient mammalian HEK293-6E cell line instead of lepidopteran cells. The remodelling was performed in a three step cloning procedure developed by the author: firstly, a tripartite promoter fragment including a CMV promoter together with a corresponding cytomegaloviral immediate early enhancer was integrated; secondly, the Epstein-Barr virus oriP for episomal replication in the EBNA-1 expressing host cell line was added to the backbone and finally an improved MCS was incorporated. The cloning and subsequent proof-of-concept in HEK293-6E and BEVS were carried out under the author's direction in collaboration with Carmen Lorenz.

Initially, a PCR-fragment of the promoter region of pTriEx (Novagen) was synthesised comprising the immediate early promoter and enhancer of the cytomegalovirus, the baculoviral late P10 and a combination of the T7 phage promoter and the lac operator. EcoRV and BamHI sites were attached by the PCR-primer design. This fragment was integrated into pFlpBtM-I upon the excision of a section containing the baculoviral ie1 promoter, the hr5 enhancer and the late baculoviral p10 promoter by ZraI-BamHI digest (Figure 3-12). However, all clones that had been identified as positives in analytical restriction digest showed a deletion of 61 bp within the intron region as confirmed by sequence analyses. Hence, this deletion had to be a result of an incorrect PCR-product due to a looping out of the highly repetitive intron sequence during elongation. Since upstream and downstream sequences had been integrated correctly, the full length intron could be incorporated by a NdeI-PacI restriction fragment of the parental pTriEx. The resulting intermediate vector was named pFlpBtM-TriEx.

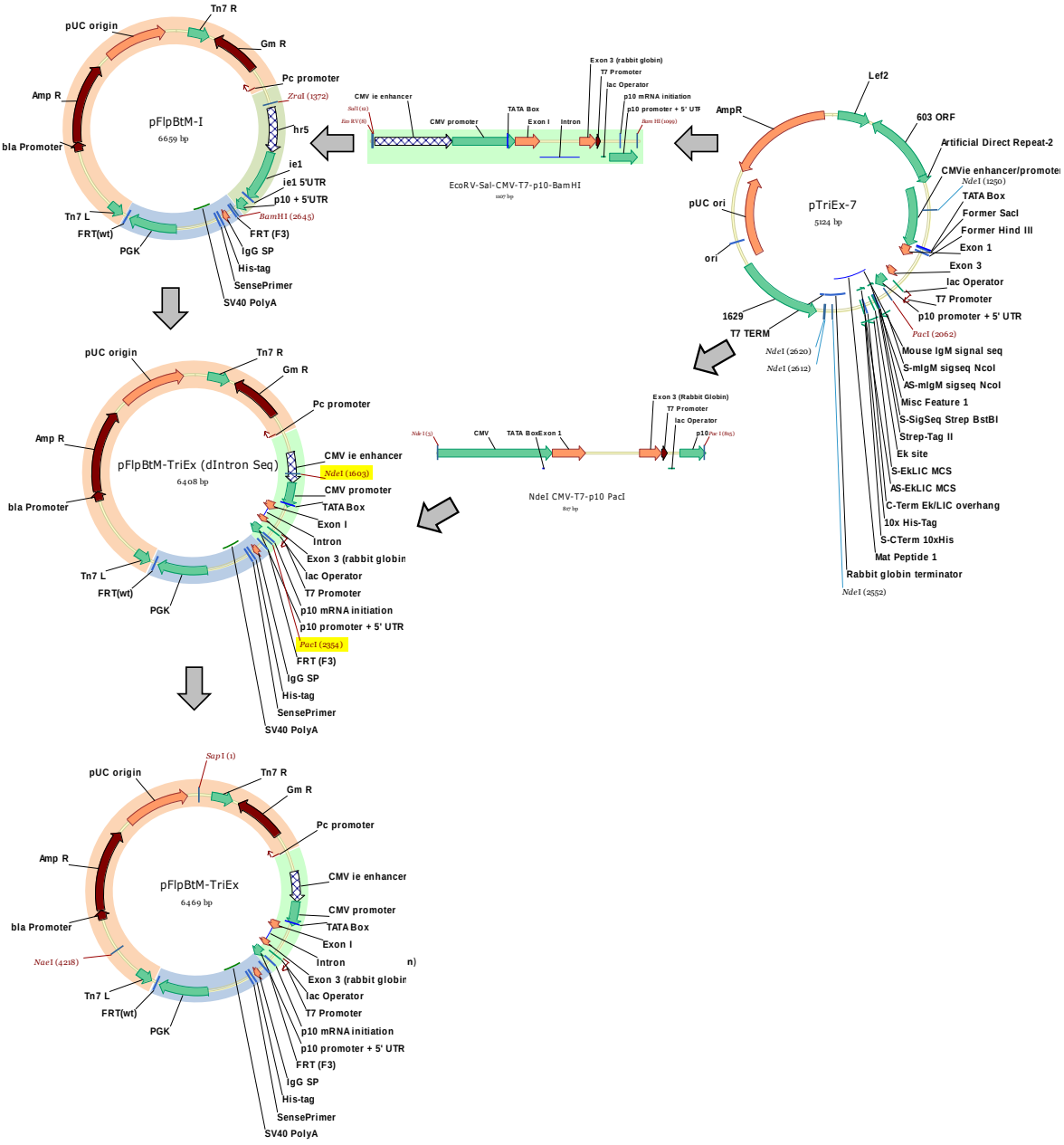


Figure 3-12: Exchange of the promoter fragment in pFlpBtM-I
The original promoter fragment containing the baculoviral ie1 promoter and the hr5 enhancer besides the late baculoviral p10 promoter was excised by ZraI-BamHI digest. A PCR-fragment of the promoter region of pTriEx-7 (Novagen) was generated with flanking EcoRV-BamHI sites. This section contains the CMV promoter for expression in mammalian cells, the p10 promoter for baculoviral expression and an inducible T7-lac promoter for prokaryotic expression. Sequencing revealed a deletion of 61 bp within the Intron which was corrected by the exchange of a NdeI-PacI restriction fragment of the right length. The intermediate product was termed pFlpBtM-TriEx.

Prior to the exchange of the backbone of pFlpBtM-TriEx by the backbone of pTT5 containing the oriP sequence, the template had to be modified since it contains NcoI and BbsI restriction sites which have to be used as unique sites in the MCS. Therefore, NcoI was eliminated by site directed mutagenesis. To remove the BbsI site in pTT5-Nco^{mut}, a

PCR-fragment of the section between XmaI-BbsI was generated and subsequently integrated into the vector. Due to the orientation of the BbsI sequence, the corresponding reverse primer was constructed to eliminate the BbsI site upon reintegration.

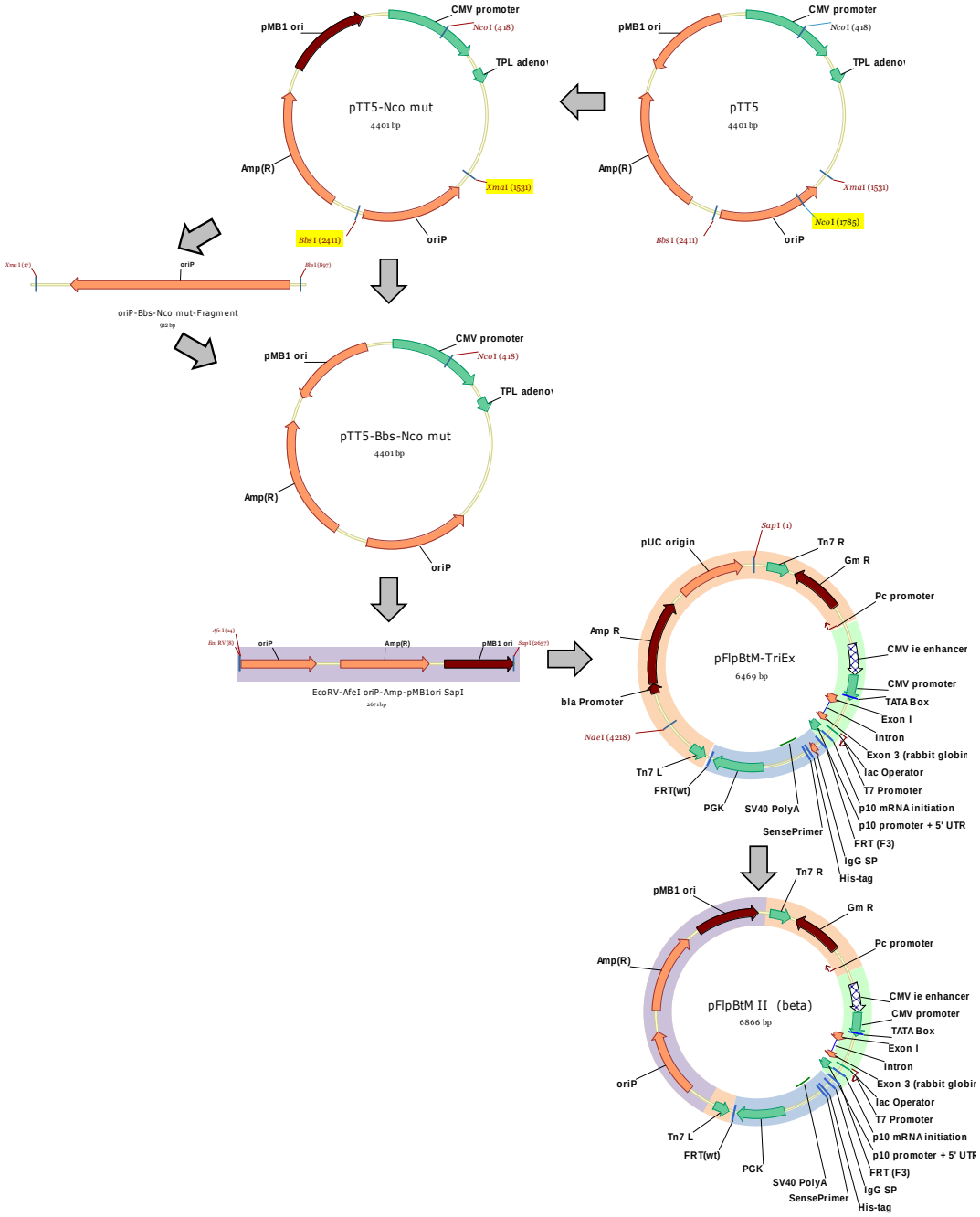


Figure 3-13: Modification of the pTT5 backbone and its integration into pFlpBtM-TriEx
NcoI in the backbone of pTT5 was mutated by site directed mutagenesis, whereas a PCR-fragment of the section between XmaI-BbsI was generated and subsequently integrated into the vector. The reverse primer was designed in a way that the BbsI site was eliminated due to its orientation. The eliminated restriction sites are highlighted in yellow. The backbone of pFlpBtM-TriEx was excised by NaeI-SapI digest and replaced by a PCR-fragment of the modified pTT5 backbone with flanking EcoRV-SapI sites. The resulting vector was named pFlpBtM-II(beta).

The resulting pTT5-Nco-Bbs^{mut} was used as a template to synthesise a PCR-fragment of the backbone sequence with flanking EcoRV and SapI sites. It was subsequently integrated into pFlpBtM-TriEx upon excision of its original backbone by NaeI-SapI digest. The resulting pFlpBtM-II(beta) (Figure 3-13, page 71) comprises all elements for its use in RMCE, BEVS and transient expression in HEK293-3E, and thus was used for the proof-of-concept studies as presented in chapter 3.4. In parallel, the final pFlpBtM-II version was generated by the integration of a remodelled RMCE cassette as shown in Figure 3-14.

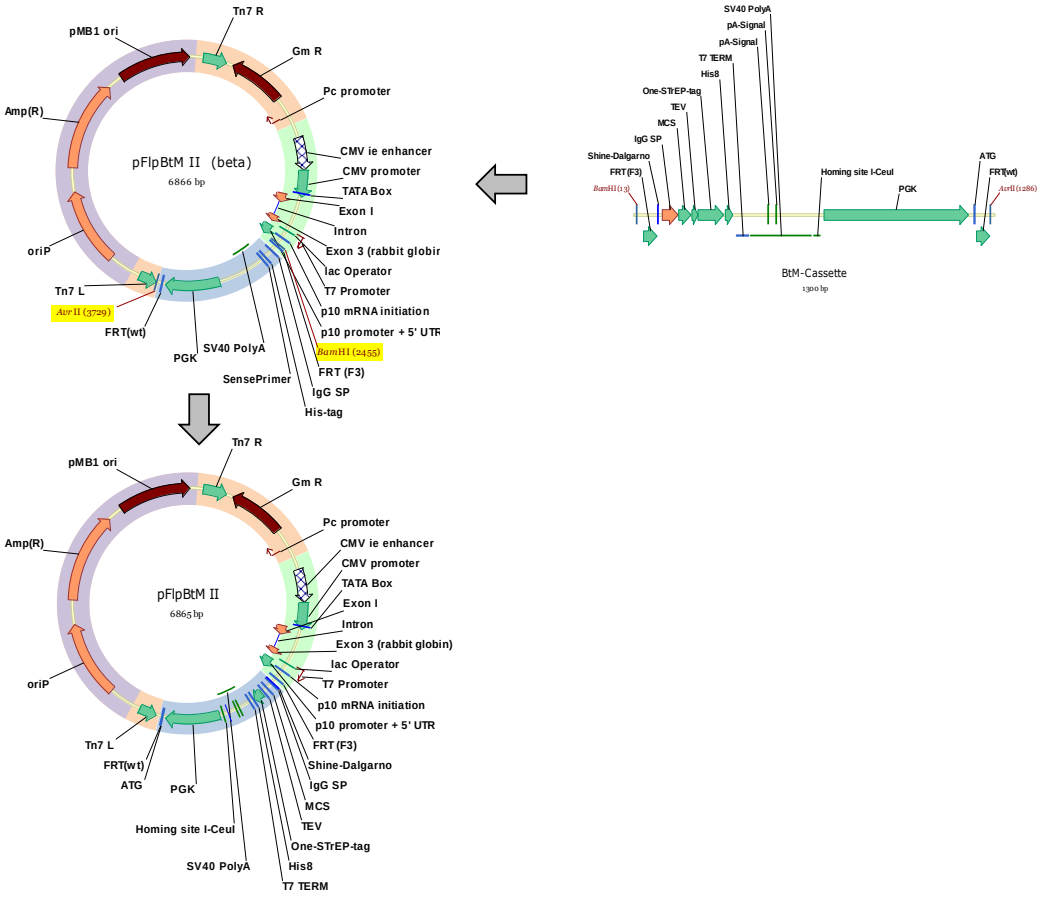


Figure 3-14: Modification of the FRT-cassette in pFlpBtM-II
The complete FRT-cassette ranging from the p10 promoter and the downstream FRTwt sequence has been synthesised to implement unique restriction sites flanking the FRT sites and the PGK promoter and to include a One-STrEP affinity tag separated from the modified MCS by a TEV protease cleavage site. Moreover, a Shine-Dalgarno sequence was incorporated between FRT$_3$ and the NcoI site necessary for expression in *E. coli*. The fragment was integrated into pFlpBtM-II(beta) by digestion and ligation via BamHI-AvrII sites (highlighted in yellow).

The synthesised sequence contained a One-STrEP affinity tag, additional restriction sites within the MCS and flanking the elements used for recombination as well as a Shine-Dalgarno sequence to enhance translation in *E. coli*. It was integrated into pFlpBtM-II(beta) upon excision of the corresponding original fragment by BamHI-AvrII. Figure 3-15 shows a schematic map of the final construct and highlights the sections derived from different parental vectors from which it was composed of.

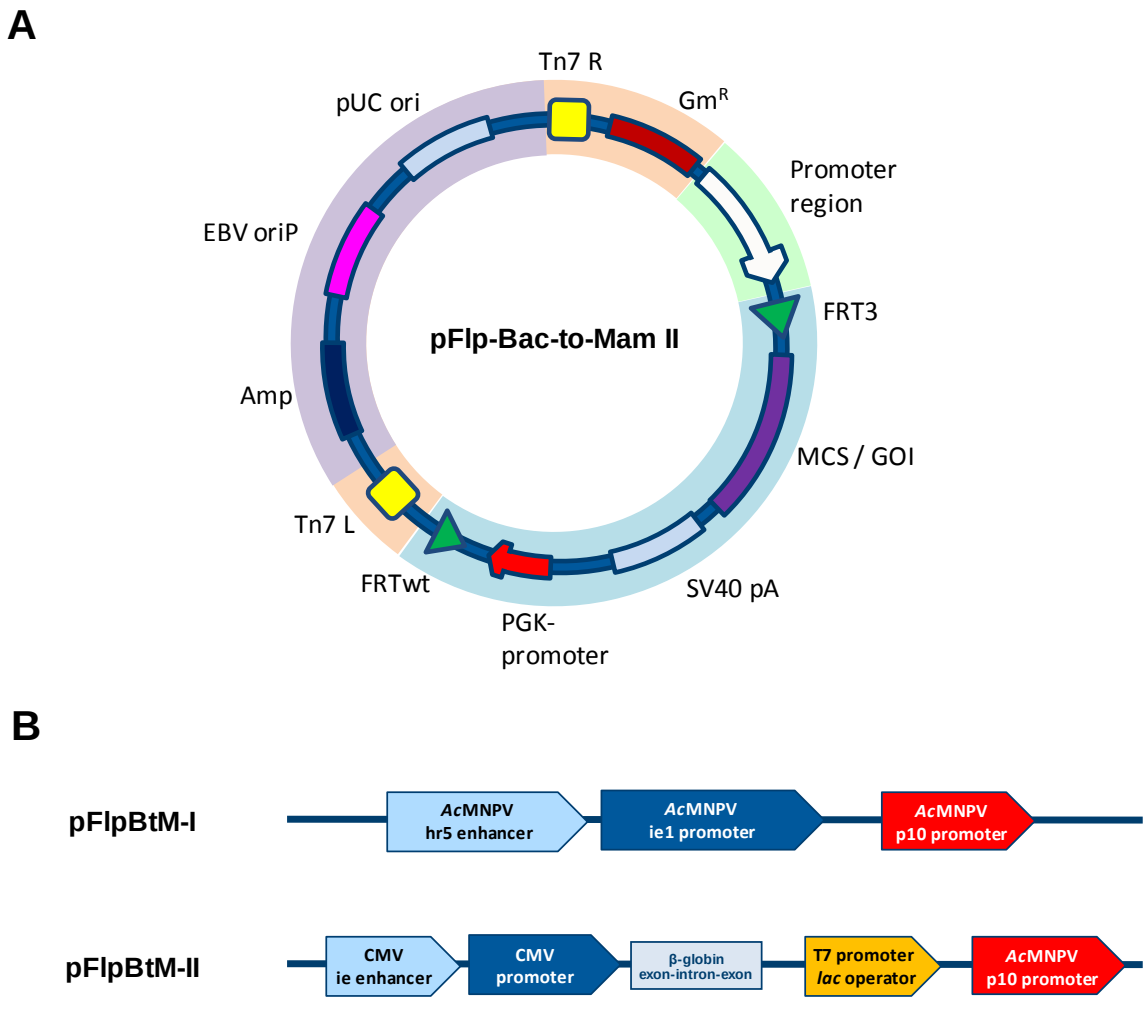


Figure 3-15: Schematic map of pFlp-Bac-to-Mam-II
A: Vector map of pFlpBtM-II with elements for recombination, expression, screening and replication. Sections originating from different parental vectors are highlighted in orange (pFastBac), blue (synthetic fragment based on pFS-sighis-PGKd), green (pTriEx-7) and purple (pTT5), respectively. **B:** Comparison of promoter regions of pFlpBtM-I and pFlpBtM-II. The immediate early *Ac*MNPV promoter ie1/hr5 and the very late P10 in pFlpBtM-I drive protein expression in transient and baculoviral expression, respectively. To enable transient expression in mammalian cells pFlpBtM-II includes the CMV promoter/enhancer instead. An additional lac-controlled T7 promoter should allow for protein expression in *E. coli*.

3.4 Evaluation of pFlpBtM-II

To proof the applicability of pFlpBtM-II in different expression systems, the fluorescent model proteins eGFP, mCherry and tdTomato were used as reporters. The vector constructs were generated as described in Figure 3-4 C on page 61 by cloning the corresponding genes from pFlpBtM-I into pFlpBtM-II(beta). The resulting constructs were fusions with a C-terminal His_6 tag without the IgG signal peptide to achieve intracellular accumulation.

3.4.1 Transient Expression in HEK293-6E

The evaluation pFlpBtM-II(beta) constructs in transient expression in HEK293-6E was performed as described in 2.6.1. Deviating from the protocol no pTTo/GFP control plasmid was added to prevent interfering fluorescence signals. Growth kinetics and fluorescence characteristics are presented in Figure 3-16. Transfection rates of more than 80 % could be achieved for eGFP and tdTomato constructs. In contrast, pFlpBtM-II(beta) driven expression of mCherry could only be detected in approximately 50 % of the cells by flow cytometry. This is an effect of the innate weak brightness of mCherry and improper excitation by the 488 nm laser in the Guava EasyCyte™. Thus, only those cells which took up a higher plasmid dose and therefore accumulate more reporter protein are detected by the cytometer. In turn, fluorescence microscopic observations revealed similar transfection rates and fluorescence intensities among all three constructs since optimal excitation and emission parameters provided by the filter sets at the microscope allow for accurate signals (Figure 3-17, page 76).

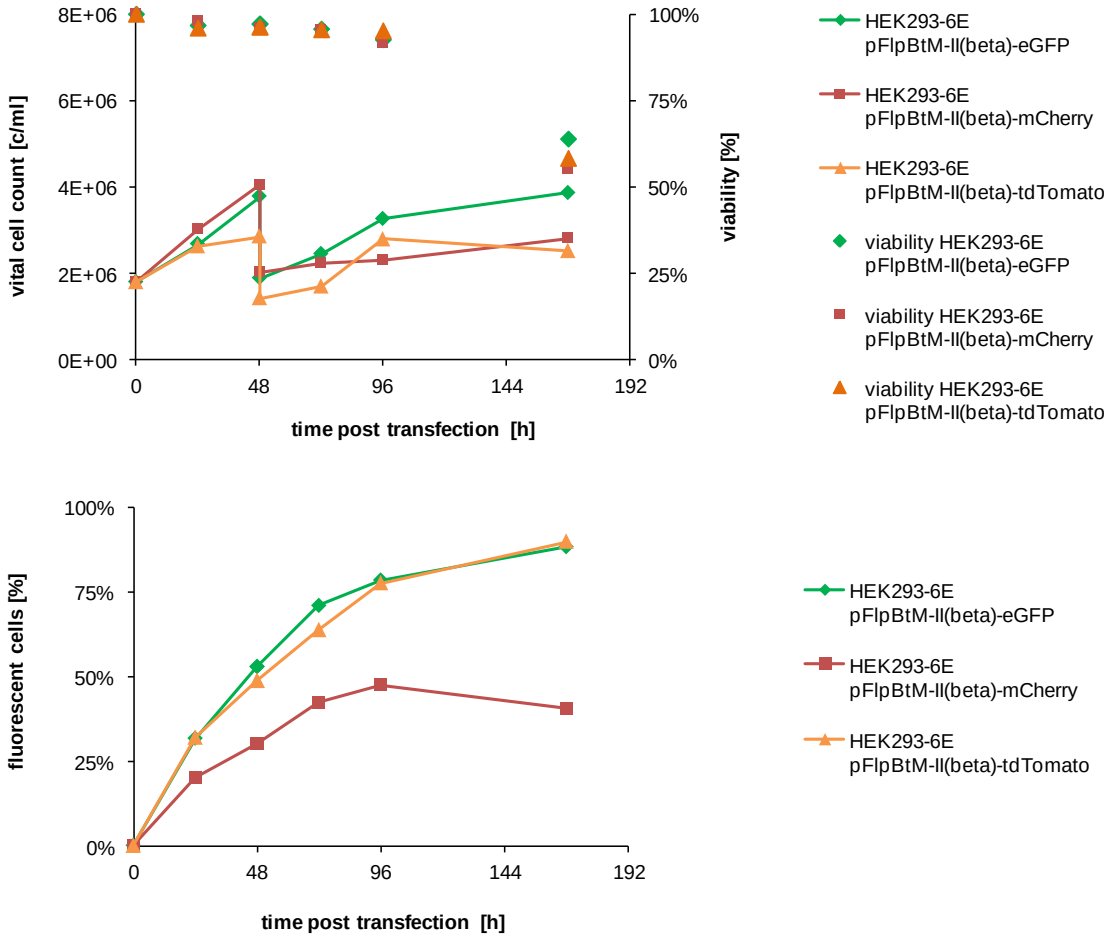


Figure 3-16: Growth kinetics and fluorescence of HEK293-6E transfected with pFlpBtM-II(beta) constructs
Three days upon seeding of 0.5 × 10^6 cells/mL HEK293-6E were transfected with complexes of 25 µg plasmid and 50 µg PEI. The culture was expanded to 50 mL 48 h post transfection and supplemented with 0.5 % tryptone. To compensate the metabolic consumption 4.5 g/L glucose were added 72 hpt. Addition of 3.75 mM valproic acid to enhance transcription rate 96 hpt caused a decrease of viability due to its cytotoxicity. Fluorescence was determined at a Guava EasyCyte flow cytometer using a 488 nm laser for extinction.

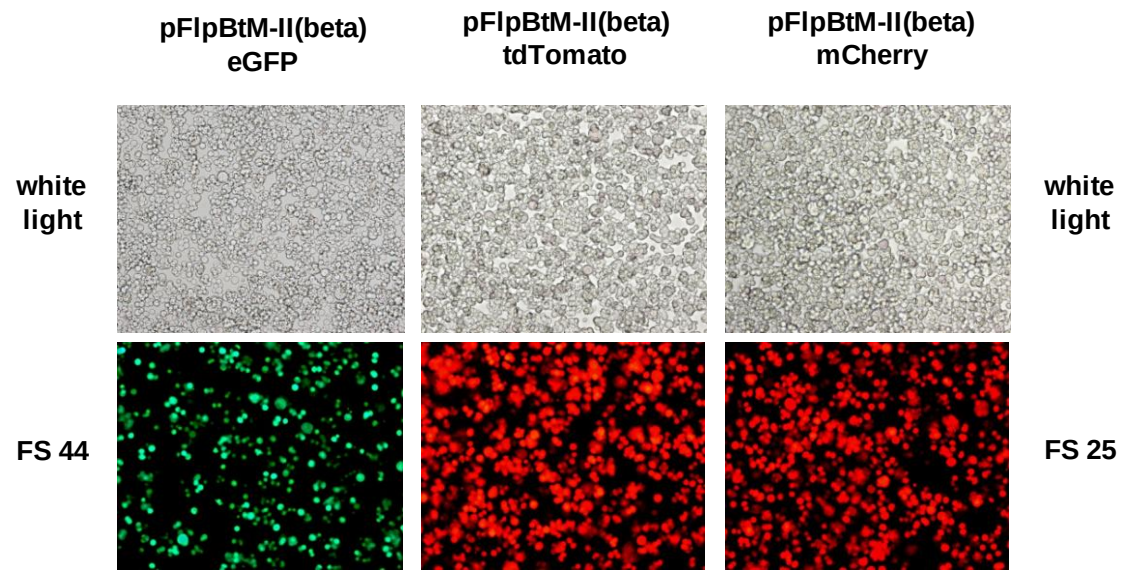


Figure 3-17: Fluorescence microscopy of HEK293-6E transfected with pFlpBtM-II(beta) constructs
Pictures were taken 96 h post transfection under white light and illuminated by a HBO50 mercury-vapor lamp using Zeiss filter sets 44 (eGFP) and 25 (mCherry/tdTomato), respectively. The exposure time for fluorescence detection was 1 s.

The overexpression of the model proteins could not only be detected by flow cytometry and fluorescence microscopy, but also was evident by stained cultures. A photo of cell pellets of HEK293-6E 96 h post transfection is shown in Figure 3-18. They exhibit the characteristic colours of mCherry, tdTomato and eGFP even in daylight without excitation by specific wavelength.

Figure 3-18: Cell pellets of HEK293-6E upon transfection with pFlpBtM-II(beta) constructs
Cell pellets of 1×10^6 cells harvested 96 h post transfection with pFlpBtM-II(beta)-mCherry (left), -tdTomato (middle), and -eGFP (right).

Likewise, the reporter proteins could be detected on coomassie stained SDS-PAGE gels of cell lysates (Figure 3-19). Significant accumulation of the model proteins can be detected approximately 48 h post transfection. Starting 96 hpt proteins also appear in the culture supernatant due to cell lysis caused by the cytotoxic effect of valproic acid.

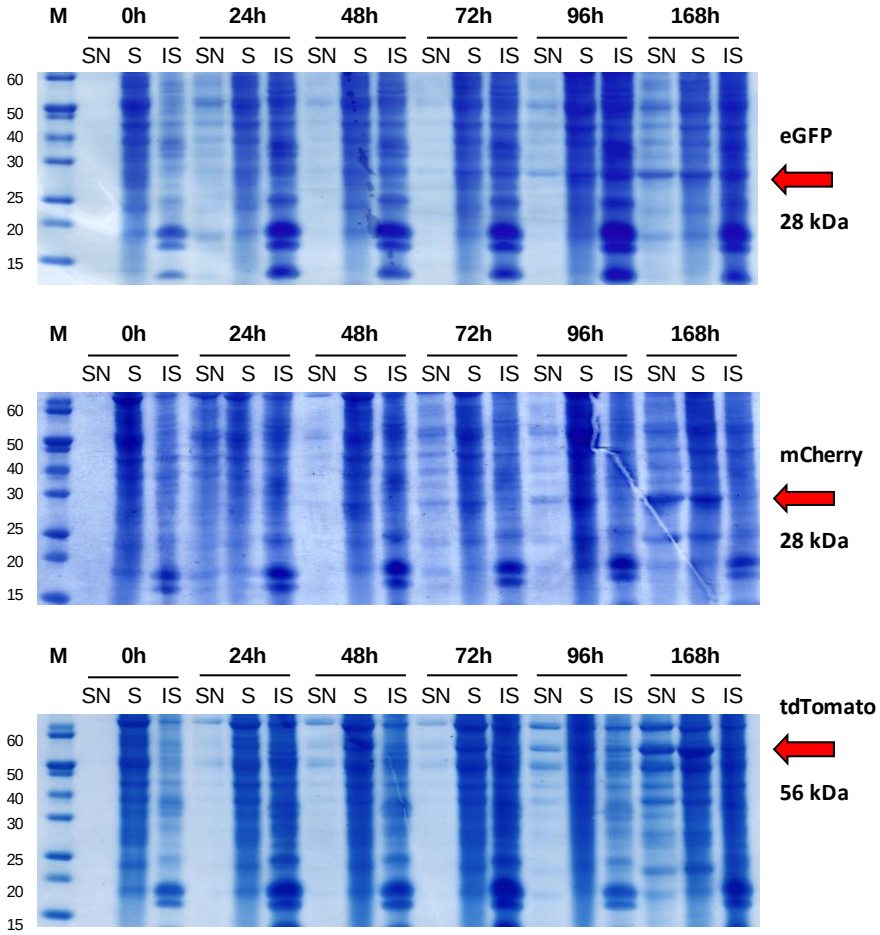


Figure 3-19: Intracellular fractions and supernatants of HEK293-6E transfected with pFlpBtM-II (beta) constructs
Samples of culture supernatant (SN) as well as cell lysates (IC) or soluble (S) and insoluble intracellular fractions from non-transfected cells (oh) and from sampling between 24 and 168 hpt were analysed. The model proteins give signals at 28 kDa (eGFP & mCherry) and 56 kDa (tdTomato) , respectively. M = PageRuler™ Protein Ladder [kDa]

To further evaluate the production of the model proteins and their quality, small scale affinity chromatography of cell lysates was performed. Upon chemical disruption of 5×10^6 cells the soluble fraction was applied to QIAGEN Ni-NTA spin columns. SDS-PAGE analyses of the IMAC fractions are presented in Figure 3-20.

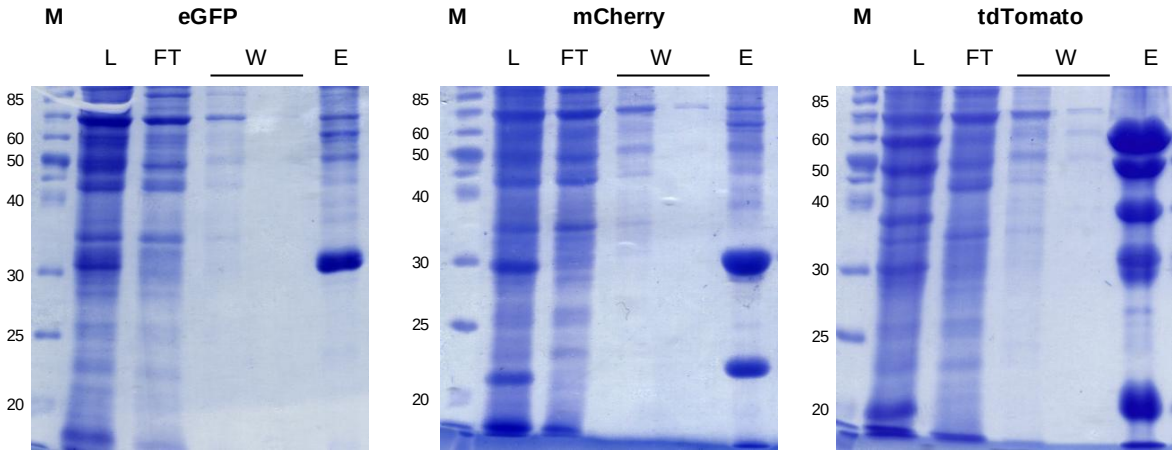


Figure 3-20: IMAC fractions of HEK293-6E lysates after transfection with pFlpBtM-II(beta) constructs
5×10^6 HEK293-6E transfected with pFlpBtM-II(beta)-eGFP (left), -mCherry (middle) and -tdTomato (right) have been harvested 196 h post transfection. Lysates were purified using QIAGEN Ni-NTA spin columns. The full length model proteins give signals at 28 kDa (eGFP & mCherry) and 56 kDa (tdTomato), respectively. M = PageRuler™ Protein Ladder [kDa], L = lysate, FT = flow through, W = wash fraction, E = eluate

A significant enrichment of the recombinant His-tagged model proteins in the eluate could be achieved. However, due to the simplicity of the spin column purification noteworthy contamination with nonspecific bound proteins with high molecular weights occurred. Notably, in the purification of mCherry and tdTomato fragments smaller than the full length proteins have also been enriched. Each of these proteins could be identified as a fragment of the corresponding model proteins by MALDI-TOF mass spectrometry (data not shown). Strikingly, a comparison of the theoretical fragment size according to the first amino acid of the fragments determined by N-terminal sequencing with the apparent size of the fragments in the gel revealed, that two fragments of each protein must also have been truncated by C-terminal degradation (indicated by asterisks in Table 3-1). Since the C-terminal His_6 tag was utilised for the purification, the tdTomato fragments could only be enriched as protein dimers via co-purification. This is not possible for the mCherry fragments, since mCherry is an obligate monomer. Therefore, those fragments had to be caused by intramolecular nicking. This was confirmed by a native PAGE where both proteins migrated as untruncated monomeric molecules (Figure 3-21). Only denaturing conditions in SDS-PAGE cause the molecules to disassemble into the observed fragments. Since the fragmentation also occurred upon expression in both insect cells and CHO Lec3.2.8.1 (cf. chapter 3.5.1) this phenomenon is protein specific. As a result, the tremendous overexpression of the model protein proved the unrestricted applicability of pFlpBtM-II for transient expression in HEK293-6E.

Table 3-1: Identification of protein fragments by N-terminal sequencing

Protein	fragment size (gel) [kDa]	# of N-terminal amino acid	calculated size [kDa]
tdTomato	~ 55	2	55
(483 aa)	~ 40	198	32*
	~ 35	165	35
	~ 30	185	34
	~ 27	247	27
	~ 22	382	19
	~ 10	247	27*
mCherry	~ 28	2	28
(243 aa)	~ 22	75	19*
	~ 10	2	28*

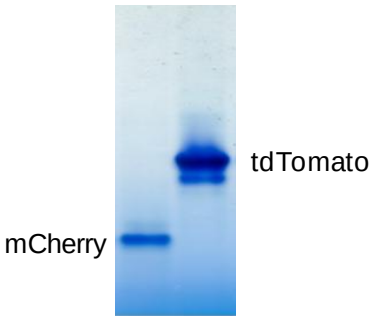


Figure 3-21: Native Page of mCherry and tdTomato after purification by IMAC
mCherry and tdTomato purified by IMAC from cell lysates of HEK293-6E were applied on a Blue Native PAGE. Under native conditions no truncated fragments of either protein occurred.

3.4.2 Generation of recombinant bacmids

In analogy to the evaluation of pFlpBtM-I in BEVS as presented in 3.2.2, recombinant baculoviruses were generated using pFlpBtM-II(beta)-eGFP, -mCherry and -tdTomato as donor vectors according to the described protocols. EMBacY derived viruses were employed for the expression of tdTomato and mCherry whereas a MultiBac bacmid was used for the expression of eGFP again. Subsequent to the screening for positive recombinants by blue-white selection and analytical PCR, Sf21 cells were transfected with isolated bacmids. Following the amplification of virus harvested from the transfection and titer determination, test expressions in Sf21 were performed using an MOI of 2. Infection kinetics were consistent with those observed in the test expressions with pFlpBtM-I (page 65). Flow cytometric measurements and fluorescence microscopy confirmed infection

rates of more than 90 % (Figure 3-22). The overexpression and intracellular accumulation of the proteins were also indicated by the characteristic staining of both the culture and cell pellets and by the presence of prominent bands on SDS-PAGE gels (Figure 3-23).

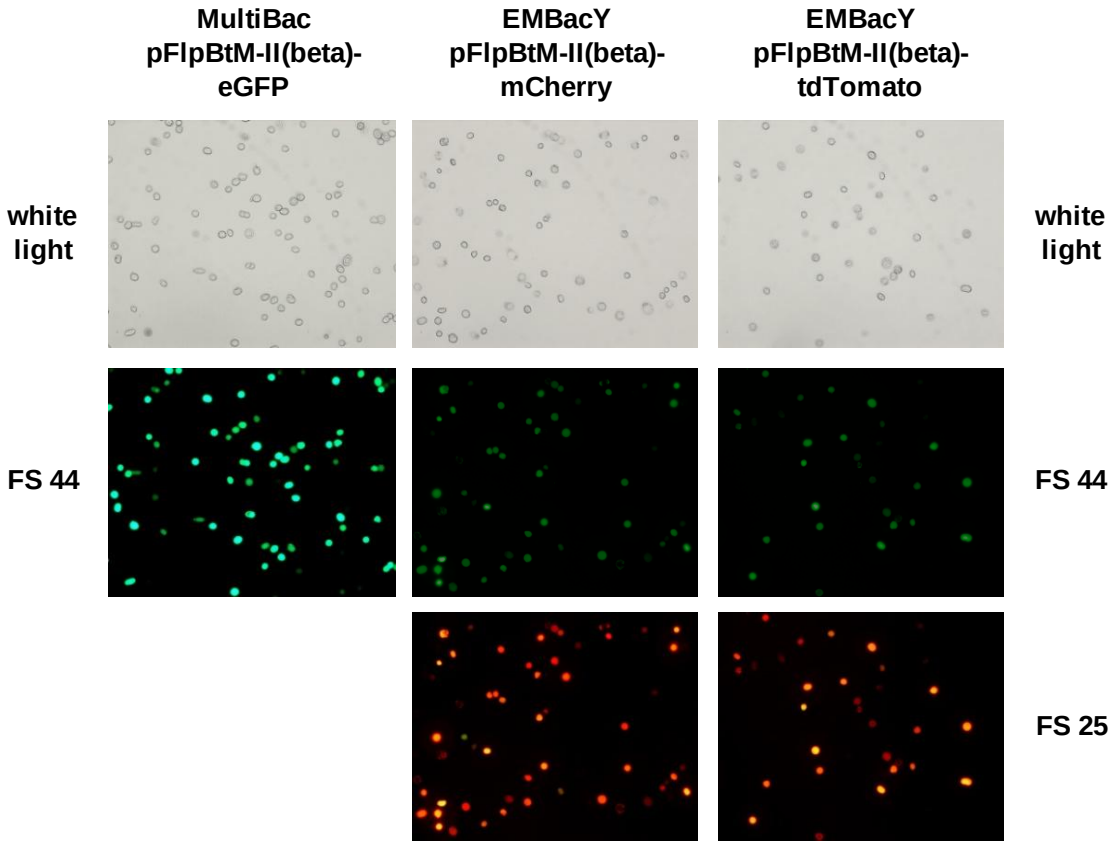


Figure 3-22: Fluorescence microscopy of Sf21 infected with pFlpBtM-II(beta) derived recombinant baculoviruses
Pictures were taken 48 h post transfection under white light and illuminated by a HBO50 mercury vapour lamp using Zeiss filter sets 44 (eGFP/YFP) and 25 (mCherry/tdTomato), respectively. The exposure time for fluorescence detection was 500 ms. Due to the specifications of the filter set and the higher quantum yield of eGFP the fluorescence of this protein appears brighter compared to YFP expressed by the EMBacY-derived viruses in this pictures.

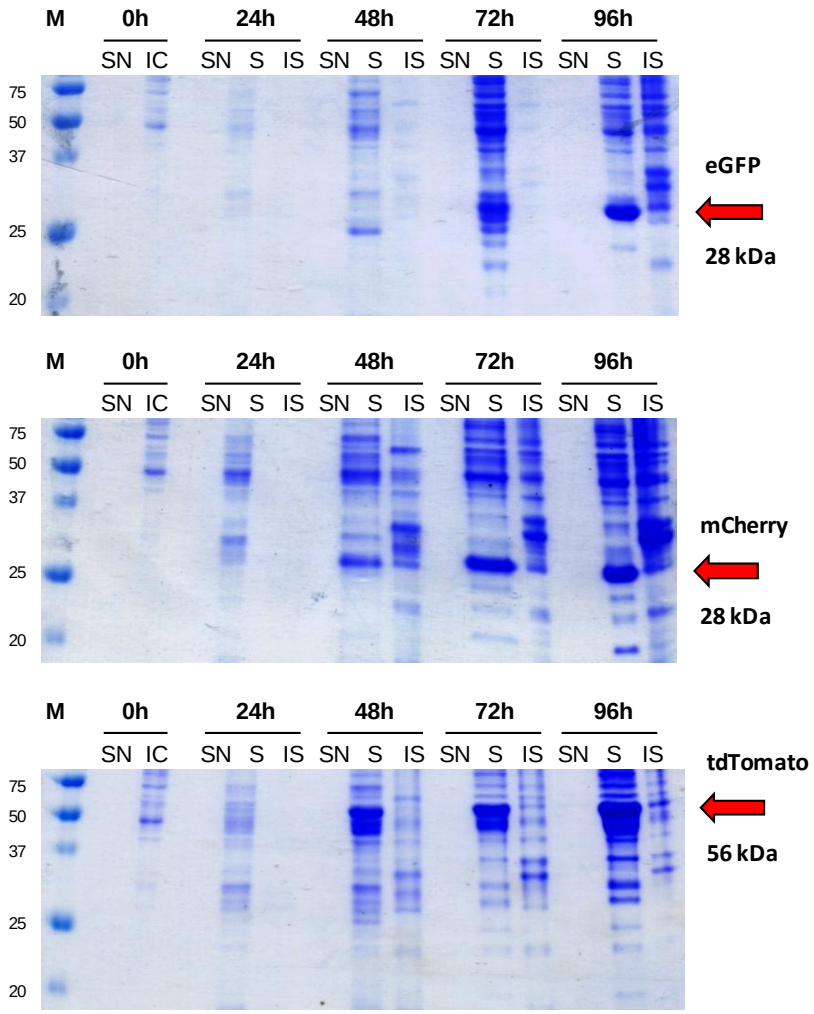


Figure 3-23: Intracellular fractions and supernatants of Sf21 infected with pFlpBtM-II(beta) derived viruses
Samples of culture supernatant (SN) as well as cell lysates (IC) or soluble (S) and insoluble intracellular fractions from non-infected cells (oh) and from sampling between 24 and 168 hpt were analysed. The model proteins give signals at 28 kDa (eGFP & mCherry) and 56 kDa (tdTomato), respectively. M = BioRad Precision Plus Protein Standard [kDa]

Western blots confirmed that the indicated bands result from the overexpression of the model proteins (data not shown). Moreover, the Western blots revealed the occurrence of the truncation fragments of mCherry and tdTomato which were observed in the purification of the proteins after expression in HEK293-6E before, thereby probating that the fragmentation is a protein specific phenomenon. These results evidently show that the improved pFlpBtM-II(beta) is likewise suitable for the use as a donor vector for protein expression in the BEVS.

3.4.3 RMCE

The modifications which were performed to generate pFlpBtM-II(beta) do not affect the section of the vector that is recombined during RMCE. Consequently, the cassette integrated into the master cell genome in both versions is identical. However, to exclude any influence of the altered vector backbone on its efficiency the applicability of pFlpBtM-II(beta) as a donor vector for RMCE has been assessed again. To investigate whether the gene length has any influence on the RMCE success tdTomato was used as a reporter gene in this experiment. As it is a dimer of DsRed, its gene is twice as large as the previously tested mCherry. However, no complications during the procedure occurred and the generation of a clonal producer cell line was accomplished within 12 weeks. Stability and homogeneity of the genomic expression was again analysed by fluorescence microscopy and flow cytometry. The corresponding microscopic pictures, dot plot and histogram are shown in Figure 3-24.

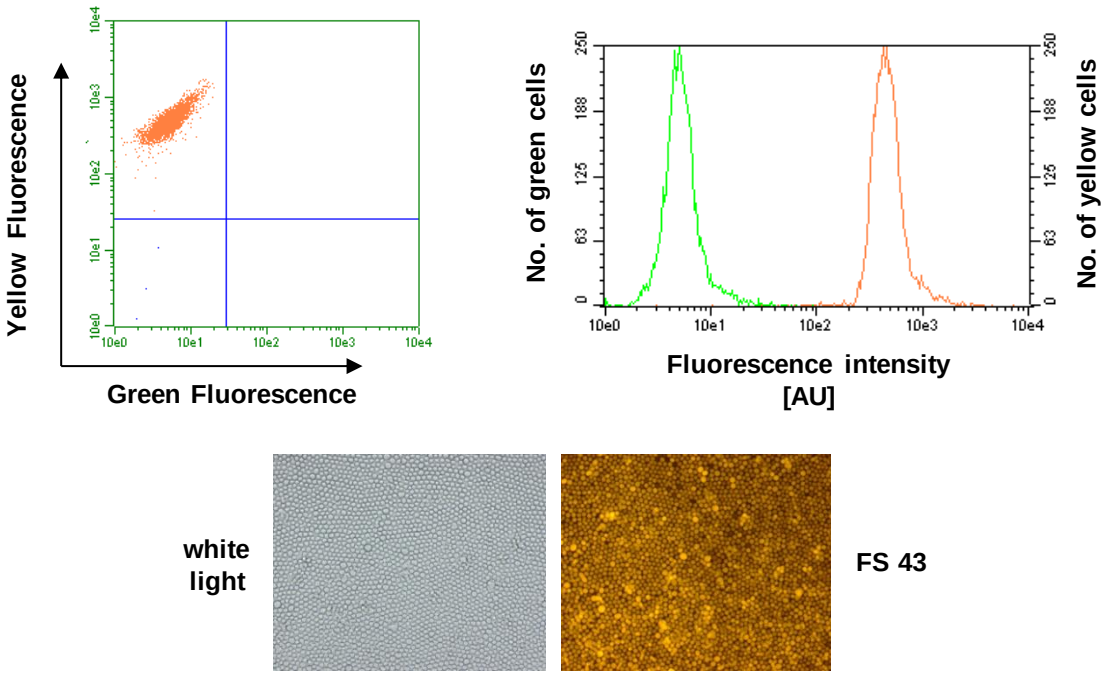


Figure 3-24: Fluorescence characteristics of the CHO Lec3.2.8.1 tdTomato producer line
Yellow fluorescence due to stable expression tdTomato in CHO Lec3.2.8.1 after cassette exchange with pFlpBtM-II(beta)-tdTomato was assessed at the Guava EasyCyte Mini cytometer and fluorescence microscopy. Upon excitation with a 488 nm laser, strong yellow fluorescence is observed whereas no green fluorescence could be detected in both dot plot (upper left) and histogram (upper right). In contrary to the previous pictures the tdTomato specific filter set Zeiss 43 HE was available to determine the expression of the reporter. The pictures were taken upon illumination by a HBO50 mercury-vapor lamp with an exposure time of 500 ms (bottom, picture taken by Bahar Baser).

3.5 Production of model proteins

Based on the successful proof-of-concept of pFlpBtM-II as a multi host donor- and expression vector for protein production in mammalian cells and the BEVS, comparative test expressions were performed to demonstrate its applicability for an evaluation of the optimal expression strategy and the most suitable expression host for different target proteins. Therefore, mCherry, single-chain Fv-hFc constructs and ECDmTLR2 have been produced in larger scale in the diverse hosts and subsequently purified by affinity chromatography to evaluate the quality and yield of the products. To ensure reproducible data, the non-stable expressions were performed in duplicates or triplicates. Additionally, pFlpBtM-II was used for an intensive screening for expressible constructs of the ECD of human TLR5, which has not been purified before.

3.5.1 mCherry

Transient productions of mCherry in HEK293-6E and viral expressions in Sf21 have been performed in 50 - 100 mL scale using pFlpBtM-II(beta)-mCherry as expression vector and corresponding bacmids derived thereof. Transfection rates of >70 % were achieved in plasmid based transient expression in HEK293-6E cells. Likewise, more than 90 % of the Sf21 cells fluoresced 72 hpi indicating proper infection with recombinant pFlpBtM-derived virus (data not shown). To assess volumetric yields of stably expressed protein accumulated in the CHO Lec3.2.8.1, the producer cell line generated by RMCE with pFlpBtM-II(beta)-mCherry was expanded to 25 mL and equally harvested. His-tagged mCherry was captured from cell lysates by Ni-NTA affinity chromatography using a Profinia™ system (BioRad) followed by a desalting step. Exemplary chromatograms of the purification upon expression in HEK293-6E and BEVS are shown in Figure 3-25.

Average volumetric yields of repeated expressions are summarised in Figure 3-26. Notable amounts of 52 mg/L and 42 mg/L could be achieved by transient expression in HEK293-6E and baculoviral expression in Sf21, respectively with low standard deviations of only 11 – 16 % within the repetitions. In contrast, stable expression yielded only 10 mg/L of intracellular accumulated mCherry.

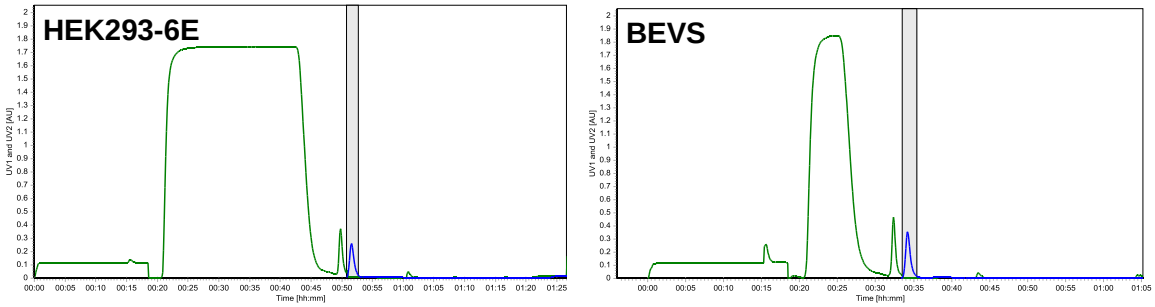


Figure 3-25: Purification of mCherry by IMAC
Cell lysates of 100 mL HEK293-6E transfected with pFlpBtM-II(beta)-mCherry (left) and Sf21 infected with EMBacY-pFlpBtM-II(beta)-mCherry (right) were applied to a BioRad Profinia system equipped with a Profinity IMAC Cartridge and a Bio-Gel P6 Desalting Cartridge. Green lines represent absorption at 280 nm directly behind the IMAC column whereas the blue lines represent absorption of the eluate upon desalting. The collected peak is highlighted in grey.

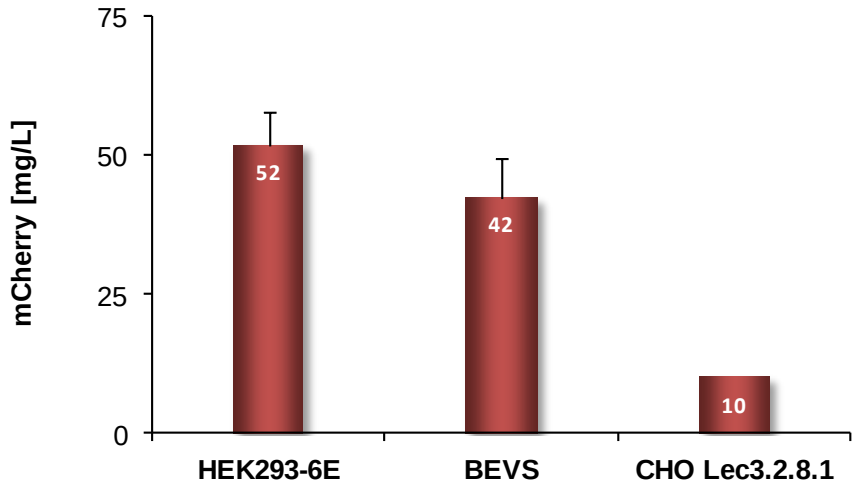


Figure 3-26: Average volumetric yields of mCherry
Average yields were calculated from the concentration of captured protein factoring the specific absorption coefficient and the original culture volume. The error bars indicate the standard deviation from triple data acquisition.

Samples of the IMAC fractions were further analysed by SDS-PAGE as shown in Figure 3-27. Minor contamination by non-specifically bound host cell protein and the characteristic mCherry "nicking-fragment" at ~22 kDa were detected in the eluate fractions of the purification from all three sources. Thus, further polishing steps including ion exchange or size exclusion would be necessary upon expression in each system to obtain crystallization grade full-length protein. The results show that in this case transient and viral expression is superior to the labour intensive generation of stable producer cell lines since higher yields of mCherry can be produced with similar purity and quality.

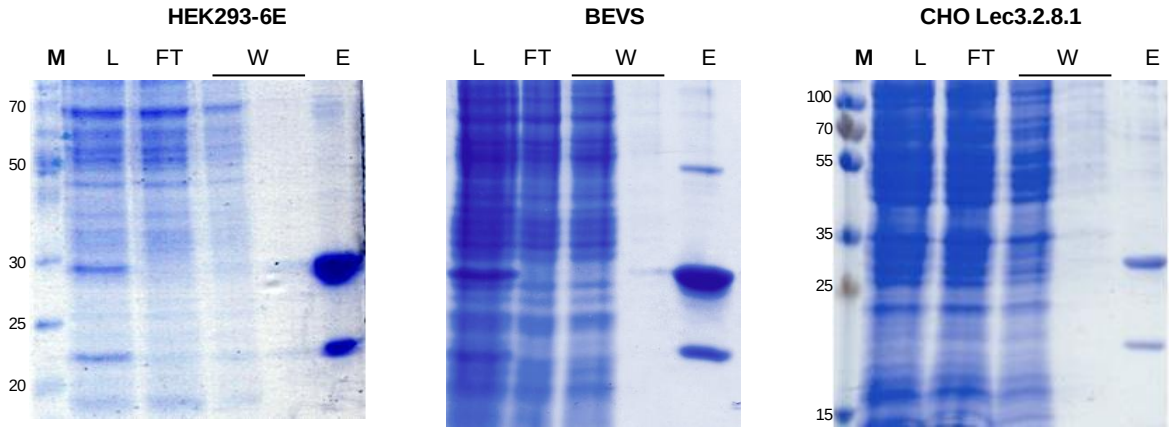


Figure 3-27: SDS-PAGE of cell lysates and IMAC fractions of mCherry purification
Samples of the cell lysates and fractions from the IMAC purification of mCherry upon transient expression in HEK293-6E (left), baculoviral expression in Sf21 (middle) and stable expression in CHO Lec3.2.8.1 (right). (M = PageRuler™ Protein Ladder [kDa], L = cell lysate, FT = flow through, W = wash fraction, E = eluate)

3.5.2 Single-chain Fv-hFc

The single-chain Fv-hFc fusion construct scFv-hIgGFc-4E3 has been provided by Thomas Schirrmann in a pCMV vector (Figure 3-28 A). It is a conventional mammalian expression vector lacking the EBV oriP for the use in HEK293-6E cells. In consequence, relatively low yields of only 2 - 4 mg/L were previously achieved in transient expression in adherent HEK293T (Menzel *et al.*, 2008). In this work the model protein was used for the benchmarking of pFlpBtM-II(beta) in comparison with both its original parental vector pCMV and pTT5, the standard expression vector for the HEK293-6E system provided by the NRCC. Therefore, a NcoI-XbaI restriction fragment of the corresponding gene was cloned into pFlpBtM-II(beta) upon digestion with NcoI-NheI. Thereby, the IgG signal peptide in the target vector was eliminated and the C-terminal His_6-tag was excluded from the ORF (Figure 3-28 B). To integrate the scFv-hIgGFc construct into pTT5 a PCR-fragment of the gene flanked by EcoRI-BamHI was generated and ligated into the target vector upon digestion with the corresponding enzymes (Figure 3-28 C). All constructs were confirmed by analytical restriction digests and sequence analyses.

Transfection rates of up to 75 % were achieved in HEK293-6E, as determined by flow cytometry and cotransfection of the pTTo/GFP control plasmid. Hence, production runs with transfection efficiencies lower than 25 % have not been included in the statistical evaluation of the protein yields.

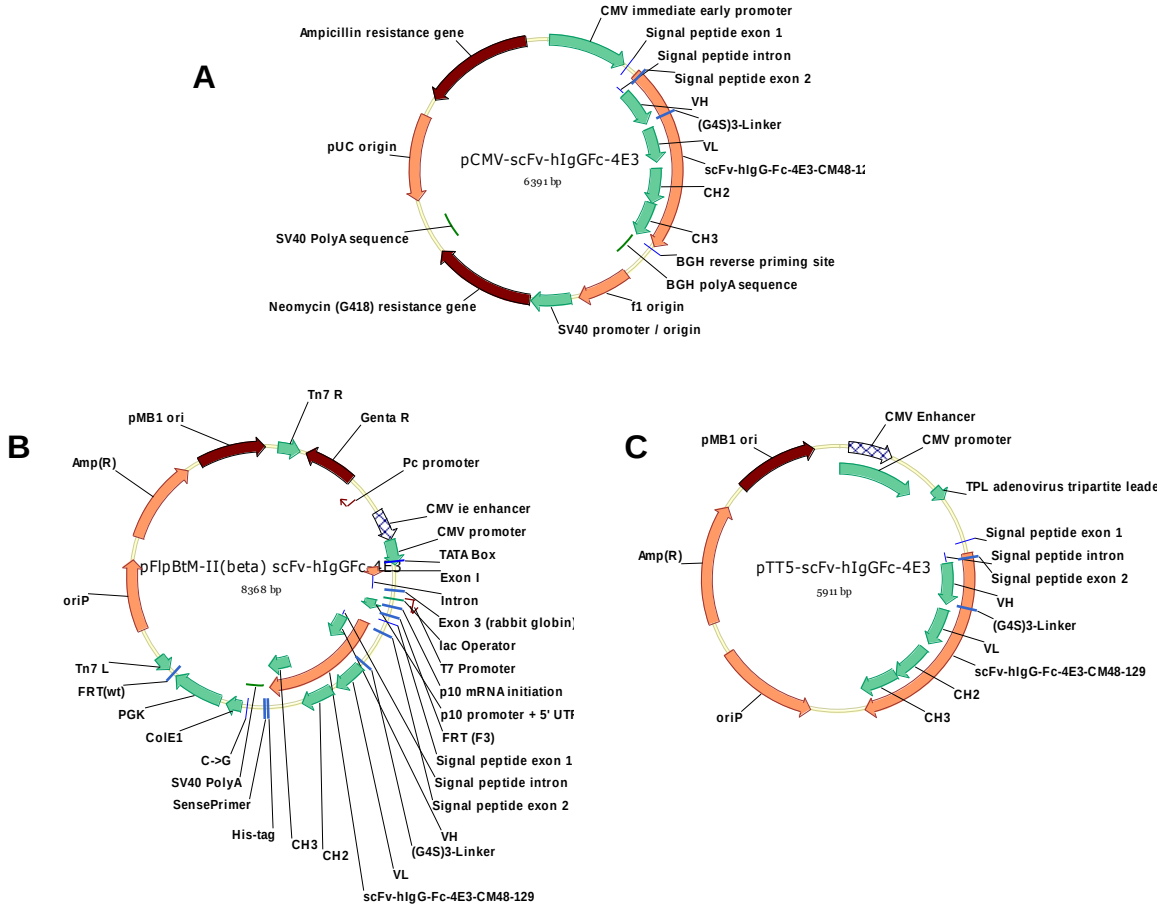


Figure 3-28: scFv-hIgG1Fc in different expression vectors
The scFv-hIgG1Fc construct was obtained in its original expression vector pCMV (A). For comparative expression in HEK293-6E cells it was cloned into pFlpBtM-II(beta) (B) and pTT5 (C). pFlpBtM-II(beta)-scFv-hIgGFc was also used to generate recombinant bacmids for baculoviral expression of the immunoglobulin.

In addition to the transient expression in HEK293-6E the scFv-construct was also produced in the BEVS to compare the achievable volumetric yields of the model protein scFv in both expression systems. Therefore, recombinant bacmids were generated using pFlpBtM-II(beta)-scFv-hIgG1Fc as donor vector. Baculoviral expression in both Sf21 and Hi5 was performed with titered 2nd generation virus using an MOI of 2 resulting in infection rates of more than 98 % in Sf21 and 76 % in Hi5, respectively (data not shown).

The model protein was captured from the culture supernatants after expression in the BEVS and HEK293-6E by Protein A affinity chromatography with the BioRad Profinia™ Purification system. Upon binding to the Bio-Scale Mini Affi-Prep Protein A Cartridge, the scFv-construct was eluted via a pH shift followed by a desalting step on a Bio-Scale Mini Bio-Gel P6 cartridge. Figure 3-29 shows typical chromatograms of the purification runs.

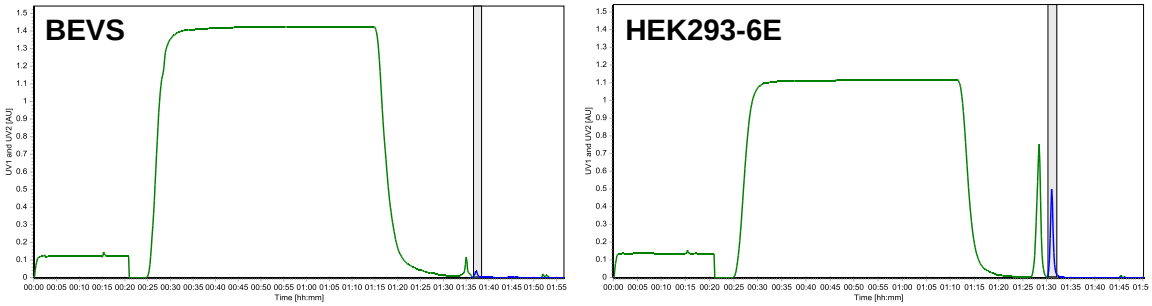


Figure 3-29: Purification of scFv-hIgGFc by ProtA affinity chromatography
Supernatants of 50 mL cultures of Sf21 infected with EMBacY-pFlpBtM-II(beta)-scFv-hIgGFc (left) and HEK293-6E transfected with pFlpBtM-II(beta)-scFv-hIgGFc (right) were applied to the BioRad Profinia system equipped with a Affi-Prep Protein A Cartridge and a Bio-Gel P6 Desalting Cartridge. Green lines represent absorption at 280 nm directly behind the IMAC column whereas the blue lines represent absorption of the eluate upon desalting. The collected peak is highlighted in grey.

Average volumetric yields of the scFv-construct from repeated expressions in both host systems are summarised in Figure 3-30. Via transient expression in HEK293-6E using pFlpBtM-II(beta) yields of ~ 50 - 120 mg/L were achieved. The average of 90 mg/L corresponds to 90 % of the amount that was captured upon expression with the optimised pTT5 vector, by which ~ 90 - 130 mg/L were produced. Expressions driven by the conventional pCMV yielded only about ~ 50 - 100 mg/L (average 69 mg/L). Although this amount achieved by PEI mediated transfection of HEK293-6E suspension cells represents a 20-fold improvement compared to the previously reported expression in adherent HEK293-T, the effect of missing episomal replication mediated by the EBV oriP on the protein yield is evident.

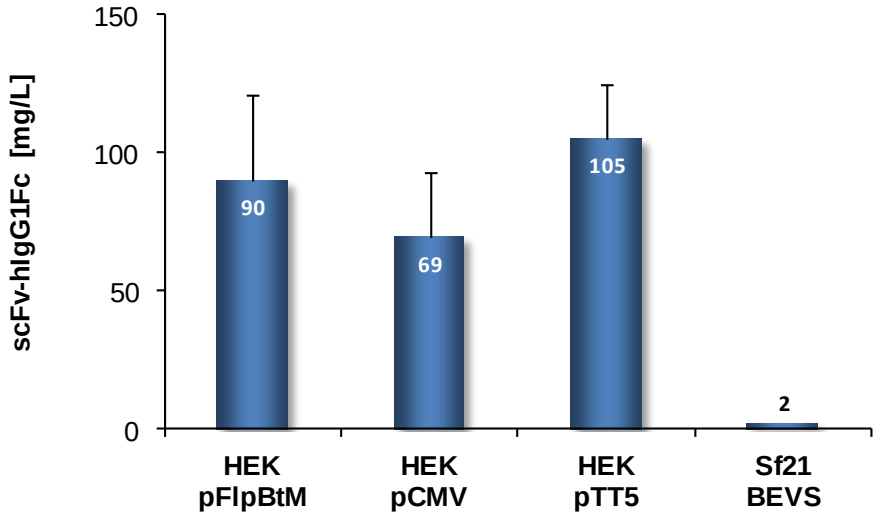


Figure 3-30: Average volumetric yields of scFv-hIgG1Fc-4E3
Average yields were calculated from the concentration of captured protein factoring the specific absorption coefficient and the original culture volume. The error bars indicate the standard deviation from triple data acquisitions.

Figure 3-31 shows an SDS-PAGE gel of culture supernatants and fractions of the chromatography of exemplary production runs.

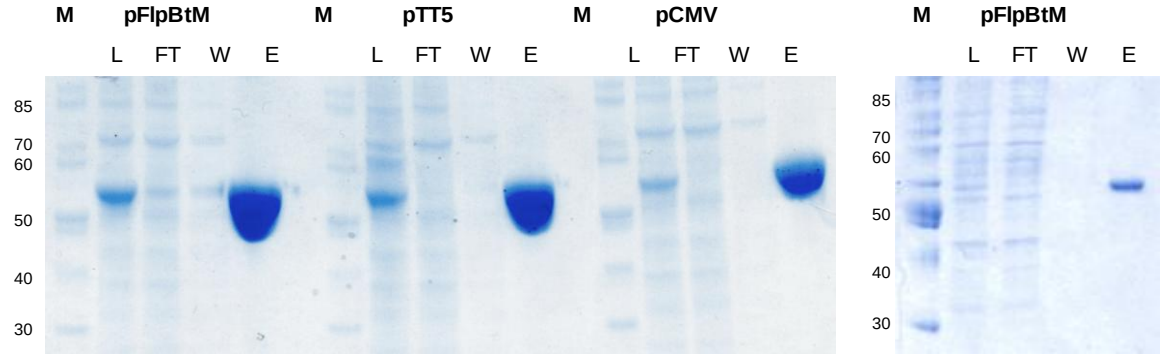


Figure 3-31: SDS-PAGE of supernatants and ProtA affinity chromatography fractions of scFv-hIgGFc purification
Samples of the culture supernatants and fractions from the affinity chromatography of scFv-hIgGFc upon transient expression in HEK293-6E (left) and baculoviral expression in Sf21 (right). (M = PageRuler™ Protein Ladder [kDa], L = cell lysate, FT = flow through, W = wash fraction, E = eluate).

Unlike the production of mCherry, which yielded similar amounts of the model protein in both HEK293-6E and the BEVS, significant differences in the yields were observed for the scFv-construct between both systems. Whereas accumulation of scFv in the culture supernatant was clearly visible in HEK293-6E expression even prior to the affinity purification, that was not the case for insect culture supernatants. Likewise, only between 1 - 3 mg/L of the model protein were captured in both Sf21 and Hi5. To determine whether improper secretion is causing the low yields of recombinant protein into the supernatant as observed with ECDmTLR2 (cf. chapter 3.2.2, page 64) Western blot analyses of cell extracts were performed (Figure 3-32). However, the results show that about 80 % of the expressed model protein are well secreted and no major intracellular accumulation of insoluble material occurred.

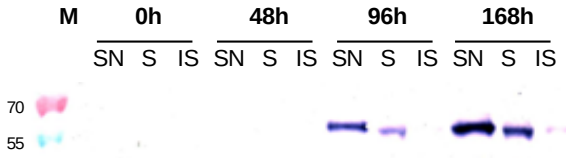


Figure 3-32: Western blot for tracing scFv-hIgGFc accumulation upon expression in Sf21
Samples of 1×10^6 cells were taken of non-infected Sf21 and 2, 4 and 7 days post infection with EMBacY-pFlpBtM-II(beta)-scFv-hIgGFc. Supernatants (SN) and intracellular soluble (S) and insoluble fractions (IS) were applied to SDS PAGE and subsequently blotted. Detection was carried out by an goat-anti-human IgG (H+L)-AP conjugate antibody (Promega #S3821). M = PageRuler™ Plus prestained Protein Ladder [kDa].

3.5.3 mTLR2

The ECD of mTLR2 has been expressed in insect cells for structural analyses before (Jin *et al.*, 2007; Kang *et al.*, 2009). However, those publications report that the expression was only possible as VLR-hybrid constructs. Notably, no expression characteristics were presented in those reports and as far as any information is given only low yields of soluble protein were achieved. In contrast, a construct comprising the first 593 amino acids of the ECD of mTLR2 has been expressed in our lab before in the BEVS using conventional pFastBac plasmids. Volumetric yields of up to 1 mg/L were achieved with this construct upon purification of heterologous protein from culture supernatants (van den Heuvel, personal communication). Yet, significant amounts of the protein remained in the cytoplasm and accumulated in the insoluble fraction (as likewise observed in the test expressions using pFlpBtM-I in chapter 3.2.2 on page 64ff). Therefore, this construct served as a model protein for an evaluation of alternative expression strategies to enhance both protein yield and quality using pFlpBtM variants.

Expression of ECDmTLR2 in the BEVS was reproduced after generating recombinant EMBacY-derived viruses using pFlpBtM-II-ECDmTLR2 as donor vector. Upon virus amplification and titer determination, expression in 100 mL culture was performed using an MOI of 2. The same expression characteristics as previously observed with a significant portion of intracellular and insoluble heterologous protein could be confirmed by Western blot analysis of samples taken 96 hpi (Figure 3-33, upper row). IMAC of diafiltrated insect cell culture supernatant led to a noticeable enrichment of secreted ECDmTLR2 so that a detection on SDS-PAGE was possible. However, significant amounts of non-specifically bound host cell proteins were found in the eluate fraction and impeded a determination of the volumetric yield of ECDmTLR2. Hence, subsequent polishing steps are necessary upon expression in BEVS.

In the transient expression of ECDmTLR2 in HEK293-6E transfection rates of more than 55 % were confirmed by cotransfection of the pTTo/GFP control plasmid. According to the baculoviral expression, Western blot analyses were performed to assess and localise the protein accumulation 96 h post transfection. The results show, that despite the non-lytic character of the system both expression and secretion of ECDmTLR2 in HEK293-6E cells was not enhanced compared to the BEVS. In fact, the protein was detected solely in the insoluble fraction (Figure 3-33, middle row, left). Even upon concentration of the

secreted ECDmTLR2 via IMAC of culture supernatants, the protein in the eluate fraction could only be detected by Western blots (Figure 3-33, middle row, right). Yet, the protein concentration in the eluate fraction was too low for an accurate determination of the yield.

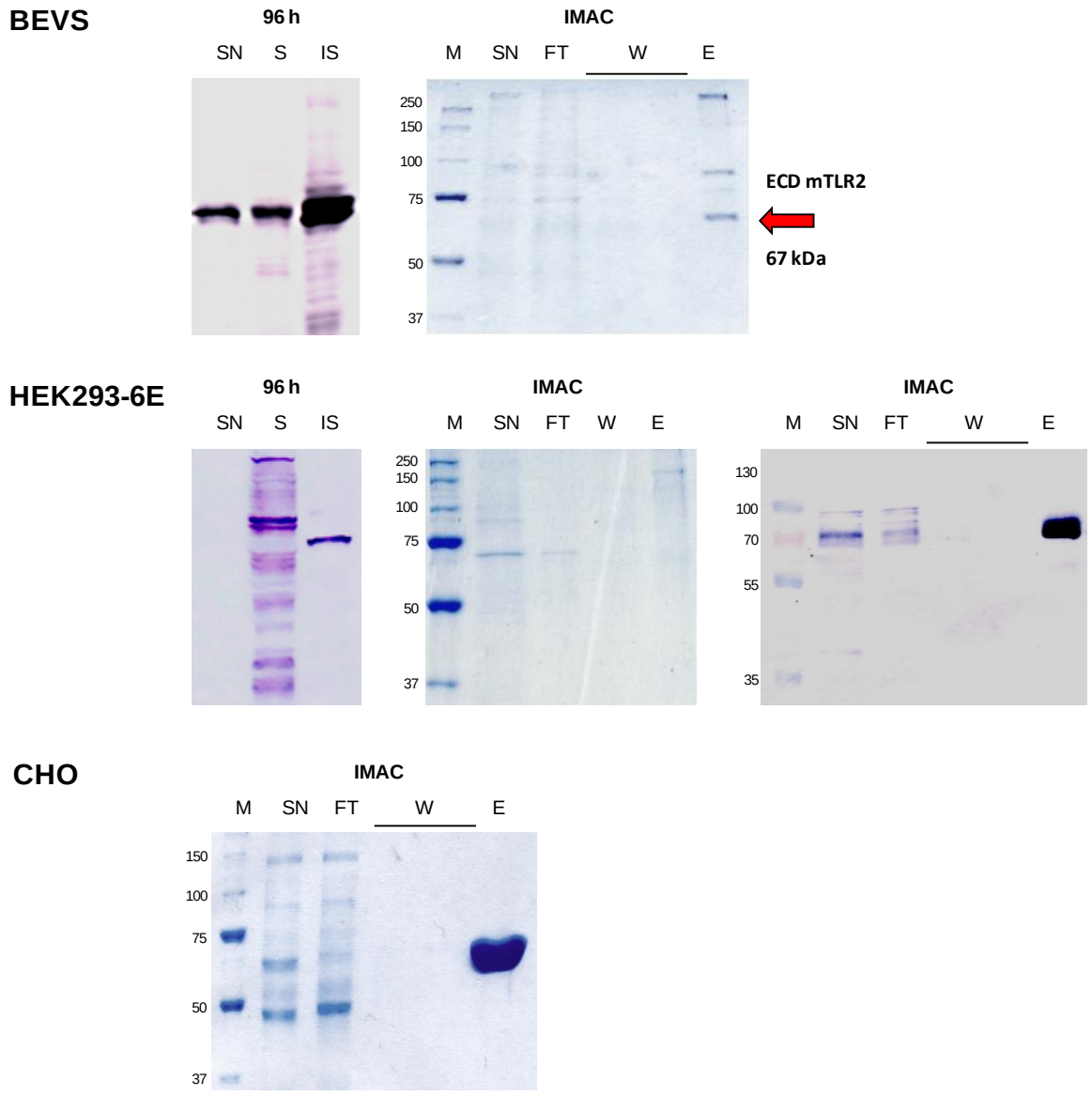


Figure 3-33: SDS-PAGE and Western blot analyses of the expression and purification of ECDmTLR2
The model protein was expressed in BEVS using an EMBacY-pFlpBtM-II-ECDmTRL2 virus, in HEK293-6E transfected with pFlpBtM-II(beta)-ECDmTRL2 and stable CHO Lec3.2.8.1 producer cell lines generated upon RMCE with pFlpBtM-I-ECDmTLR2. The upper row shows a Western blot of the supernatant (SN) and soluble (S) and insoluble (IS) fraction of 1×10^6 Sf21 96 hpi (left) and a SDS-PAGE gel of the IMAC fractions of dialyzed supernatant upon harvest. Corresponding samples of the transient expression in HEK293-6E are shown in the middle row including a Western blot of the IMAC fractions. The lower row shows the SDS-PAGE of samples of the IMAC fractions of concentrated culture supernatant upon fermentation of CHO Lec3.2.8.1 producer cell lines. Western blots were incubated with a His-Tag primary antibody and a goat-anti-mouse IgG (H+L)-AP conjugated secondary antibody. FT = flow through, W = wash, E = eluate, M = Precision Plus (Biorad) or PageRuler Plus Prestained (Fermentas) protein standard [kDa]

Whether stable expression of ECDmTLR2 in CHO improves both protein quality and yield was assessed upon the generation of a clonal producer cell line by RMCE with pFlpBtM-I-ECDmTLR2 as donor vector. Since the yield in genomic expression was expected to be low, the cultivation was performed in continuous fermentation in a bioreactor as described in 2.6.3 to accumulate significant amounts of the target protein. 35 L harvested culture supernatant were dialyzed and concentrated 25-fold by diafiltration. A 230 mL aliquot of this concentrate was used for the subsequent purification of secreted ECDmTLR2 by IMAC by which an absolute amount of ~ 1 mg ECDmTLR2 could be obtained in the eluate (Figure 3-33 bottom row). This corresponds to a volumetric yield 0.2 mg/L based on the original volume of harvested perfusion supernatant. Moreover, a 5-fold higher yield of more than 1 mg/L was achieved in 1 L batch cultivations. These amounts are not only comparable with those reached in preceding large scale purifications of ECDmTLR2 expressed in the BEVS, the stable production n CHO cells also provides the target protein in a constantly high quality necessary for subsequent structural and biochemical analyses. The multiparallel expression evaluation using pFlpBtM thereby enabled the identification of stable expression as the most suitable technique for the production of the particular challenging ECDmTLR2 target.

3.5.4 Screening for expressible constructs of ECDhTLR5 variants

In parallel to the work on the pFlpBtM project the expression and purification of the ECD of human TLR5 for structural analyses was pursued. While the X-ray structures of TLR1, TLR2, TLR3, TLR4 and TLR6 have been already solved, TLR5 still remains a promising target for structural biology. At the beginning of the work, a construct comprising the N-terminal amino acid 1 - 640 based on database prediction of the domain border was provided. The construct was generated in conventional pFastBac plasmids with C-terminal His_8 and human-Fc tags for detection and purification. It was expressed in the BEVS both with its authentic signal peptide and alternatively using the gp67 secretion signal of *Spodoptera frugiperda*. However, similar to the expression of ECDmTLR2, the protein was poorly secreted and accumulated as insoluble material in the intracellular fraction upon baculoviral expression in both Sf21 and Hi5. The deletion of the 23.6 kDa hFc-tag which might have inhibited the proper secretion in insect cells did not provide any improvements in solubility (Figure 3-34).

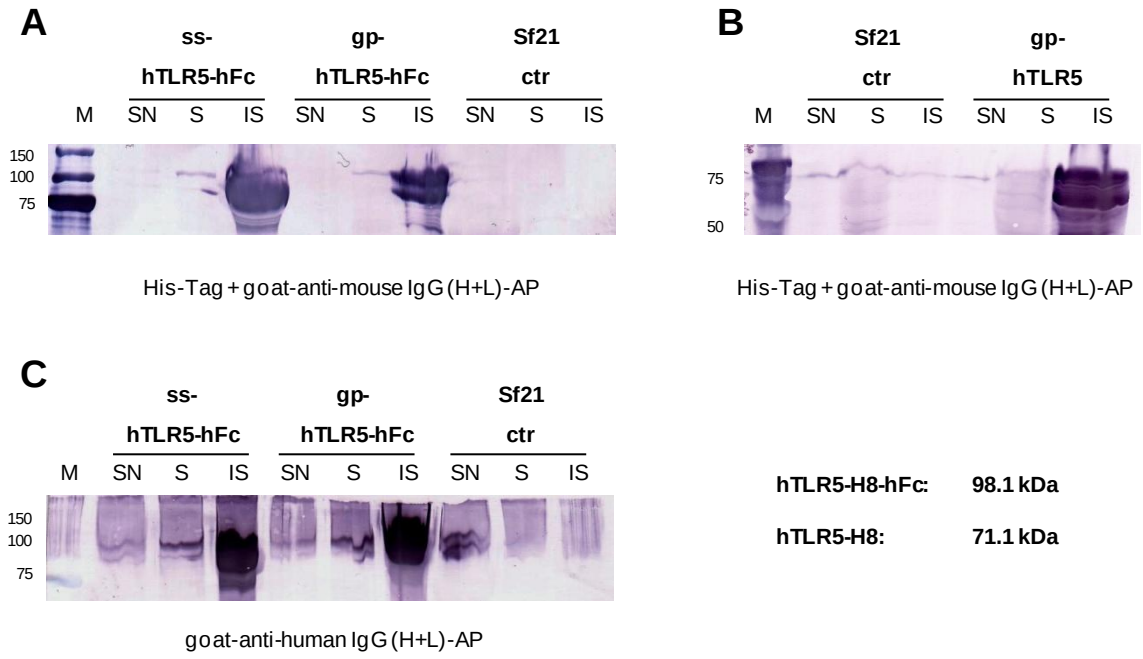
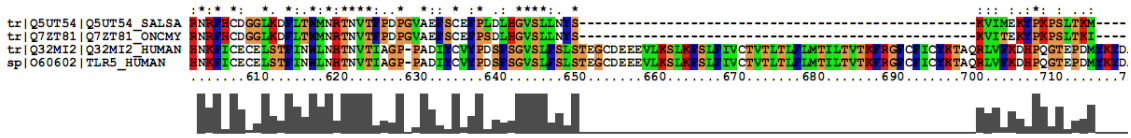


Figure 3-34: Western blots of ECDhTLR5 expression in BEVS
ECDhTLR5 was expressed as a C-terminal His_8 fusion protein in the BEVS with or without a c-terminal human Fc-tag to analyse its influence on the secretion of the protein. Different constructs with its authentic signal peptide (ss) or the gp67 (gp) secretion signal were tested. From each expression samples of 10^6 cells were taken 96 hpi and the supernatants (SN) as well as soluble (S) and insoluble (IS) intracellular fractions were analysed by Western blots using a His-Tag antibody and a secondary goat-anti-mouse IgG (H+L)-AP conjugate (A & B). Constructs with the hFc-tag were additionally incubated with a goat-anti-human IgG (H+L)-AP conjugate (C).

As the length of the original insoluble constructs based solely on putative database prediction of the domain border, a novel ECDhTLR5 construct with an alternative C-terminus was designed to achieve soluble expression and secretion. An alignment of the sequence of the full length human TLR5 and a soluble homolog of the ECD of TLR5 from the rainbow trout (Tsujita *et al.*, 2006) revealed a section of 50 amino acids at the domain border which is missing in the fish orthologs. In contrary, a 15 amino acid long C-terminal fragment behind this gap shows significant homology again (Figure 3-35 A). Based on this analysis, a chimeric variant was designed including the first 621 amino acids of ECDhTLR5 followed by the 19 amino acid rainbow-trout C-terminus. An alignment of this construct termed ECDhTLR5-rtC and the original variant is shown in Figure 3-35 B.

A



B

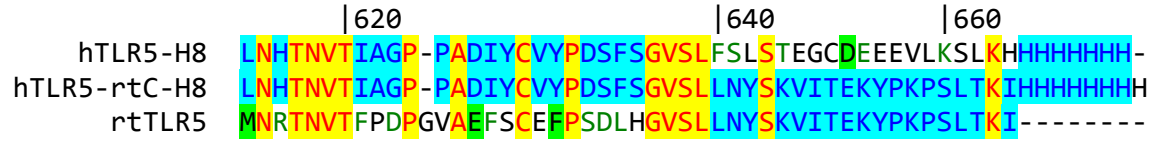


Figure 3-35: Sequence alignment of soluble rainbow trout TLR5 with human TLR5 and the chimeric construct
A: human TLR5 is aligned with the soluble variants of the rainbow trout rtTLR5 (*Oncorhynchus mykiss*, "oncmy") and the atlantic salmon (*Salmo salar*, "salsa"). The sequence identity of the ECD of human TLR5 and the soluble rtTLR5 is 27 % with 37.1 % positives. **B:** Based on the findings of the alignment, a chimeric construct (middle row) comprising the first 621 aa of human TLR5 (upper row) and the 19 aa rtTLR5 C-terminus (lower row) was generated.

Since the multipurpose vector pFlpBtM was not available at this stage the new construct had to be cloned into different conventional expression vectors by a series of cloning steps. Firstly, a synthetic 638 bp fragment containing parts of the ECDhTLR5 and the novel rtC-terminus flanked by XbaI and NotI was integrated into a pIEX vector harbouring the original construct upon excision of a corresponding wildtype segment. The resulting pIEX-sshTLR5-rtC-H8 was utilised for transient test expression in Sf21. Additionally, the chimeric hTLR5-H8-rtC construct was integrated as PCR-fragments with necessary flanking restriction sites into the vectors pTT and pFastBac for transient expression in HEK293-6E and the generation of recombinant MultiBac bacmids for baculoviral expression, respectively. However, transient expression of the construct in both Sf21 and HEK293-6E did not result in any detectable amounts of the protein. Upon baculoviral expression Western blot signals were again only observed in the insoluble intracellular fraction (data not shown).

As the expression of the chimeric construct did not enhance the solubility of the target protein either, a completely new strategy had to be employed to generate constructs which enable the production of soluble protein. Therefore, a structure model of the domain was predicted to identify secondary structure elements to design additional truncated constructs Figure 3-36.

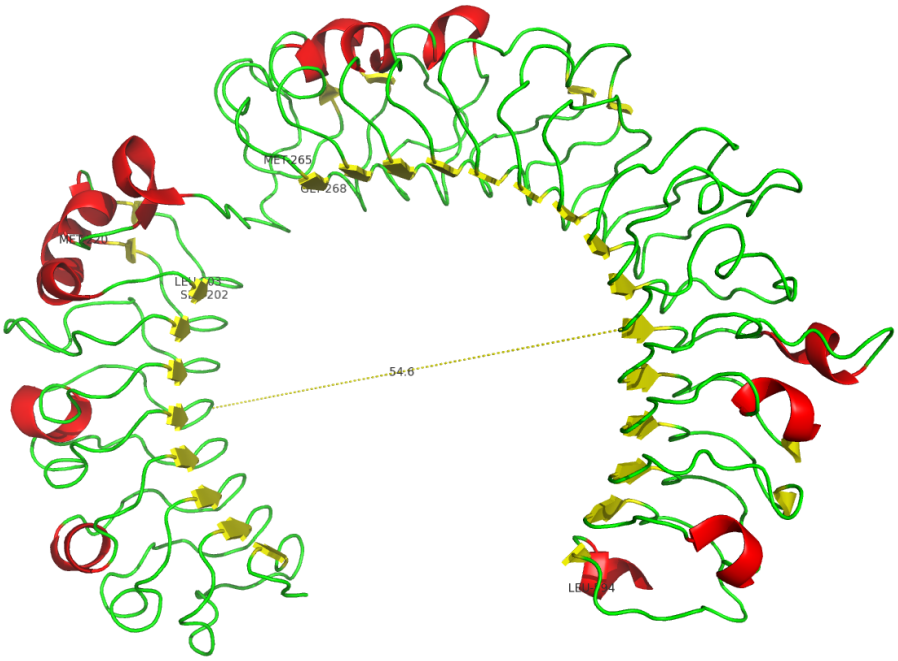


Figure 3-36: Predicted structure of the ECD of human TLR5
Structure prediction of ECDhTLR5 was performed using the CPHmodels server via remote homology modeling. The ECD of human TLR3 (PDB ID: 1ziw_A, identity 23.41 %) served as a template. The final model was kindly created by Joachim Reichelt using the PyMOL software. The diameter of the concave structure is indicated by the dashed line [Å].

Ordered structure prediction was possible only to amino acid 591. Consequently, a truncated construct comprising amino acids 1 – 591 was planned. Moreover, the model exhibits a significant gap within the parallel beta sheets typically formed by hydrophobic amino acids on the concave surface of LRR proteins between M220 and M265. Here, a disordered loop indicates a flexible region within the otherwise predominantly rigid solenoid structure. Remarkably, this section is part of the flagellin binding region which was proposed by functional studies in a 228 aa large central region between amino acids 174 - 402 (Andersen-Nissen *et al.*, 2007). Since this flexible region might inhibit proper protein folding, an N-terminal truncated construct beginning directly behind this section at M265 was designed as well.

Expression and subsequent crystallization of the ECDs of TLR1, TLR2, TLR3, TLR4 and TLR6, had been successfully performed by means of the LRR hybrid technology. Consequently, this method was also used in this work to generate novel constructs of ECDhTLR5. Therefore, 5 LRR hybrids with the C-terminal portion of hagfish variable

lymphocyte receptor VLRB6 comprising amino acids 133 - 200 were planned according to annotated LRR motifs (Matsushima *et al.*, 2007). The junctions were designed within the LRR 18-21 as shown in Figure 3-37 resulting in constructs comprising amino acids 1-480, 1-509, 1-533, 1-555 and 1-576. Furthermore, a short N-terminal fragment ending directly before LRR7 at amino acid 195 and thus upstream of the disordered region identified in the structure model was generated as well. This construct was also linked to the VLRB6-C-terminus but without generating a LRR-hybrid.

```
  1   MGDHLDLLLG VVLMAGPVFG IPSCSFDGRI AFYRFCNLTQ VPQVLNTTER
 51   LLLSFNYIRT VTASSFPFLE QLQLLELGSQ YTPLTIDKEA FRNLPNLRIL
101   DLGSSKIYFL HPDAFQGLFH LFELRLYFCG LSDAVLKDGY FRNLKALTRL
151   DLSKNQIRSL YLHPSFGKLN SLKSIDFSSN QIFLVCEHEL EPLQGKTLSF
201   FSLAANSLYS RVSVDWGKCM NPFRNMVLEI LDVSGNGWTV DITGNFSNAI
251   SKSQAFSLIL AHHIMGAGFG FHNIKDPDQN TFAGLARSSV RHLDLSHGFV
301   FSLNSRVFET LKDLKVLNLA YNKINKIADE AFYGLDNLQV LNLSYNLLGE
351   LYSSNFYGLP KVAYIDLQKN HIAIIQDQTF KFLEKLQTLD LRDNALTTIH
401   FIPSIPDIFL SGNKLVTLPK INLTANLIHL SENRLENLDI LYFLLRVPHL
451   QILILNQNRF SSCSGDQTPS ENPSLEQLFL GENMLQLAWE TELCWDVFEG
501   LSHLQVLYLN HNYLNSLPPG VFSHLTALRG LSLNSNRLTV LSHNDLPANL
551   EILDISRNQL LAPNPDVFVS LSVLDITHNK FICECELSTF INWLNHTNVT
601   IAGPPADIYC VYPDSFSGVS LFSLSTEGCD EEEVLKSLKH HHHHHHH
```

Figure 3-37: Sequence of the ECD of human TLR5
LRR motifs (LxxLxLxxNxL) are indicated in red, in which "L" usually is leucine (L), isoleucine (I), valine (V) or phenylalanine (F) and "N" is asparagine (N), threonine (T), serine (S) or cysteine (C) and "x" is any amino acid. In those LRR motifs used for the VLR-hybrids, the amino acids exchanged by the linkage are marked in yellow. The signal peptide is highlighted in blue.

A streamlined parallel construction of the mutants was facilitated by the preceding generation of an entry vector. It contained a synthesised VLRB6 fragment flanked by unique restriction sites and downstream One-STrEP-His_8 sequences separated by a TEV protease cleavage site for improved detection and purification. Since the construction and evaluation of pFlpBtM-II was still ongoing at this stage a pTT5 backbone was used for this purpose as shown in Figure 3-38. PCR-products of the ECDhTLR5 truncations were generated and integrated with technical assistance by Claudia Wylegalla. The fragments were flanked by BamHI-NheI restriction sites allowing for the integration in the linearised recipient upstream of the VLR fragment by BamHI-NheI digest or alternatively by replacing it upon digestion with BamHI-XbaI.

As soon as pFlpBtM-II(beta) was available all constructs were transferred into the multi-host expression vector as NcoI-MluI restriction fragments. Schematic illustrations of the resulting 8 new constructs as well as the preceding variants are summarised in Figure 3-39.

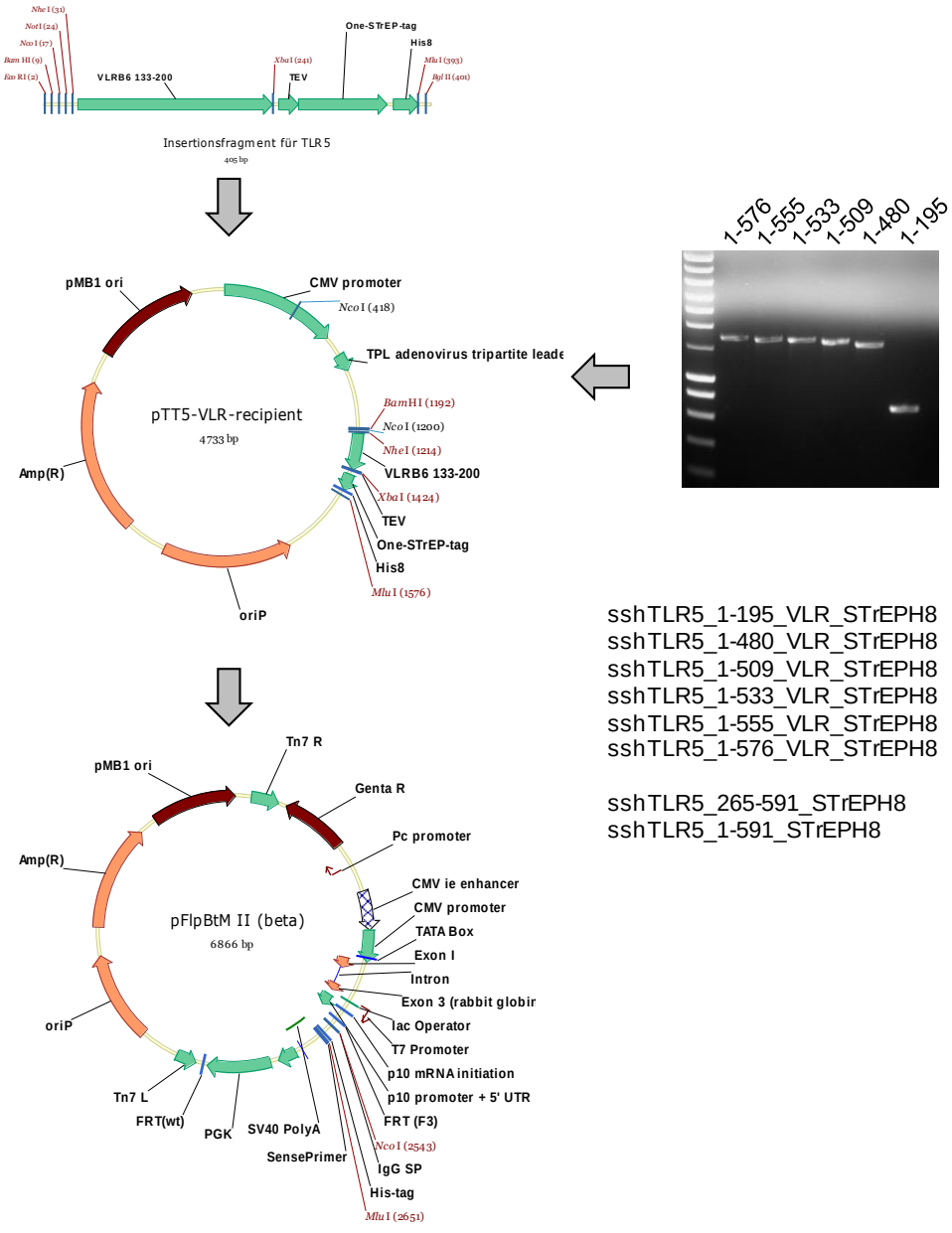


Figure 3-38: Construction VLR hybrids and truncations of ECDhTLR5
A fragment of the C-terminal amino acids 133 – 200 of hagfish VLR6 fused with a TEV-STrEP-H8 tag was synthesised and incorporated into a pTT plasmid. PCR-fragments of different truncated variants of ECDhTLR5 were generated and integrated as LRR-hybrid fusions to the VLR (upon BamHI-NheI digest) or replacing the VLR domain (upon BamHI-XbaI digest), respectively. As soon as the pFlpBtM-II(beta) was available, all constructs were transferred into the new multipurpose expression vector as NcoI-XhoI fragments.

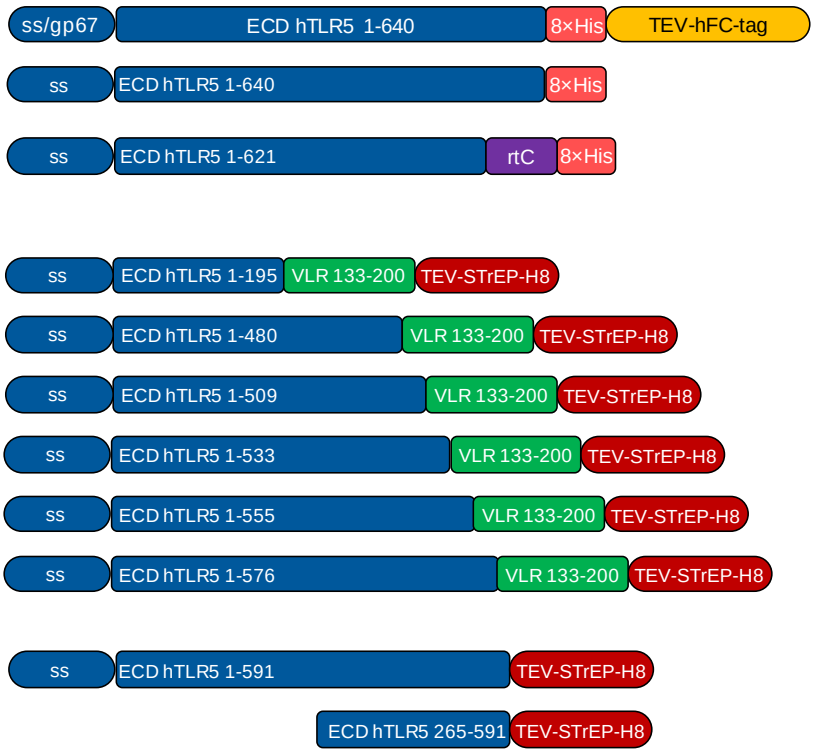


Figure 3-39: Schematic view of all 12 ECDhTLR5 variants expressed in this work
Initially, constructs comprising the amino acids 1 - 640 of ECDhTLR5 were expressed with a C-terminal hFc-tag and either its authentic signal peptide (ss) or the *S. frugiperda* gp67 secretion signal. To enhance solubility, a variant without the hFc-tag and a chimeric protein consisting of aa 1 - 621 and the C-terminal portion of rainbow trout soluble TLR5 (rtC) were constructed. Finally, a series of truncated and VLR fusion constructs were designed according to structure modelling and LRR motif predictions. Those 8 constructs were generated in an entry vector and transferred into pFlpBtM-II(beta) to screen for soluble expression and subsequent large scale expression.

The initial screening for soluble constructs was carried out by transient transfection of HEK293-6E. Average transfection rates of duplicate experiments with each construct were between 15 - 40 %. To detect expression of the variants and localise their accumulation, Western blots of supernatants and cell lysates were performed using both His-Tag and *Strep*MAB primary antibodies, respectively. As observed in transient expression of ECDmTLR2 and the previously tested ECDmTLR5 constructs, only weak signals were obtained. Especially with the His-Tag antibody a significant background of unspecific binding due to long incubations occurred (Figure 3-40). Moreover, most of the constructs predominantly yielded in insoluble material and were not secreted at all.

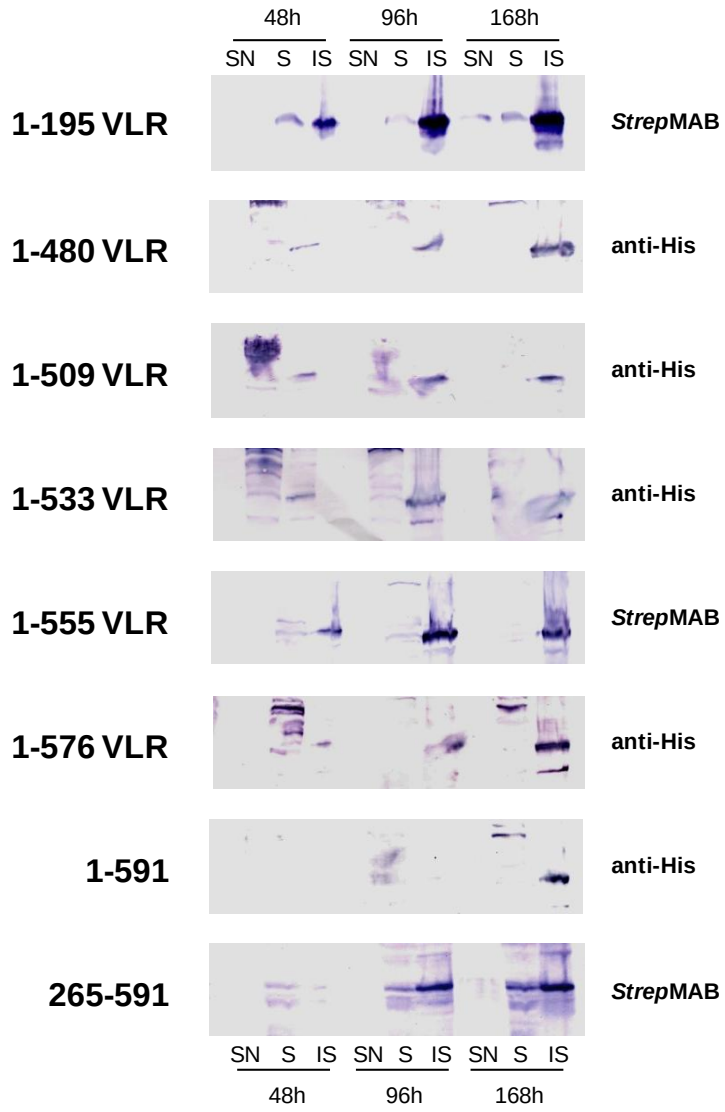


Figure 3-40: Western blots to detect expression of ECDhTLR5-VLR variants in transfected HEK293-6E
Supernatants (SN) and cell lysates (soluble: S, insoluble: IS) of 10^6 HEK293-6E cells transfected with pFlpBtM-II(beta) harbouring different ECDhTLR5 variants were taken 48, 96 and 168 hpt and analysed by Western blots using a His-Tag or *Strep*MAB antibody and a secondary goat-anti-mouse IgG (H+L)-AP conjugate.

Only the short N-terminal fragment hTLR5_1-195_VLR and the mutant hTLR5_265-591 without signal peptide or VLR fusion were detected in the supernatants or in the soluble fractions, respectively. In addition hTLR5_1-509_VLR and hTLR5_1-555_VLR exhibited significant overall expression and were slightly detectable in the soluble fraction, as well. Based on these observations, those four variants were selected to be further tested in baculoviral expression. The generation of recombinant EMBacY bacmids could be carried out directly with the pFlpBtM-II(beta) constructs. A 50 mL culture with 1 $\times 10^6$ Sf21 cells per mL were infected with 10 % 2nd generation virus and incubated for 96 h thereby achieving infection rates of approximately 90 % (data not shown).

Again, protein expression was assessed by Western blots of culture supernatants and intracellular fractions. The results are shown in Figure 3-41.

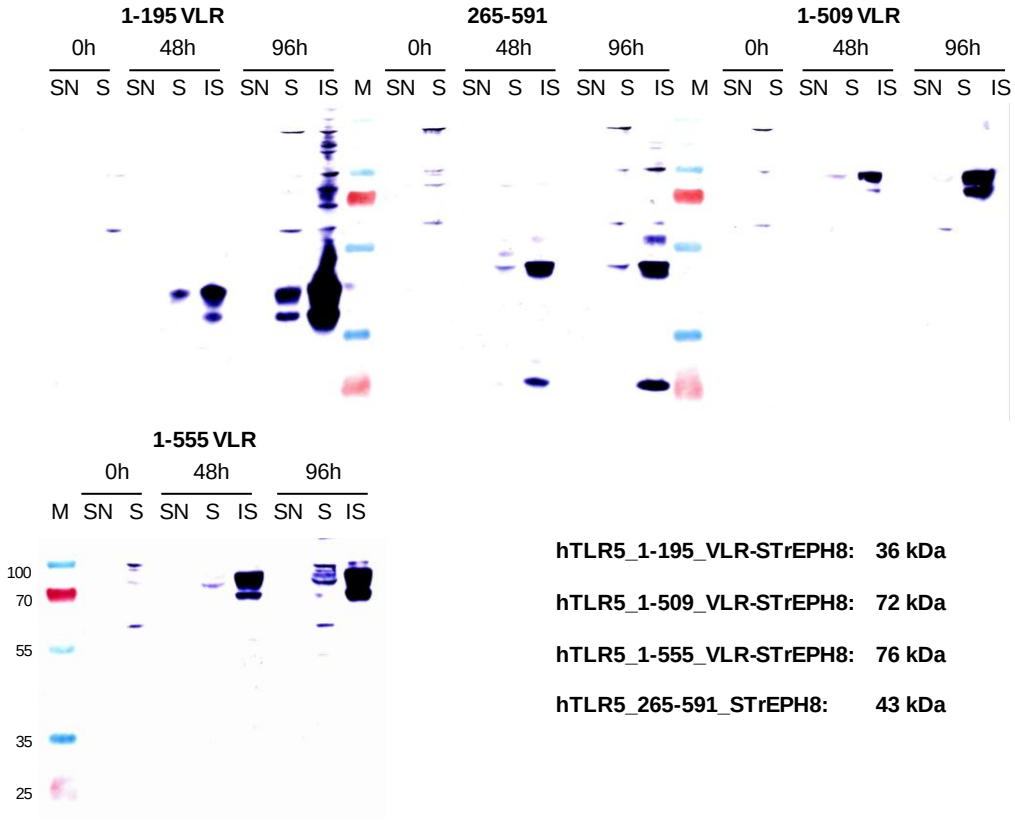


Figure 3-41: Western blots to detect baculoviral expression of ECDhTLR5-VLR variants in Sf21
Four ECDhTLR5 variants that exhibited soluble expression in HEK293-6E cells were produced in the BEVS. From each expression samples of 10^6 cells were taken 48 and 96 hpi as well as from non-transfected cells. The supernatants (SN) soluble (S) and insoluble (IS) intracellular fractions were analysed by Western blots using a His-Tag antibody and a secondary goat-anti-mouse IgG (H+L)-AP conjugate (M = PageRuler Plus Prestained [kDa]).

The Western blots of the baculoviral expression confirmed the results of the production in HEK293-6E. All constructs yielded predominantly insoluble material and only the short constructs 1-195_VLR-STrEPH8 and 265-591_STrEPH8 could be detected in the soluble fraction. Notably, in contrast to the transient expression in HEK293-6E, no secreted hTLR5_1-195_VLR-STrEPH8 was found in the supernatant. However, considerably more soluble material could be produced with the new truncated and VLR hybrid constructs compared to the preceding experiments with the full length ECD or rtC-fusions. Though, initial efforts that were made to capture intracellular accumulated soluble material in the late phase of this work have not been successful.

3.6 Engineered BEVS host cell lines

In addition to the multi-host donor and expression vector pFlpBtM-II, a BEVS host cell line was generated for stable expression of complex partners or auxiliary proteins enhancing protein folding and stability. For stable integration of target genes into the BEVS host cell the RMCE system was chosen, as it facilitates fast and robust genomic integration and furthermore would enable a compatible system to the existing CHO Lec3.2.8.1 master cell line. Thereby, an integrated expression platform for fast and versatile protein production in optimised cell lines should be provided. Likewise, fluorescence-activated cell sorting should be used for the isolation of clonal cell lines.

3.6.1 Design and construction of the tagging vector

The tagging vector design was based on pEF-FS-EGFP-dneo, which was previously used for the generation of CHO-RMCE master cell lines in our group (Wilke *et al.*, 2011). It contains the heterospecific sites FRT3 and FRTwt and is thus compatible with the pFlpBtM donor vectors. Additionally, pEF-FS-EGFP-dneo contains an ATG-deficient aminoglycoside phosphotransferase gene (Δneo) downstream of the FRT cassette used as selection-trap to screen for positive RMCE events. The ΔneoR gene is complemented and activated upon correct cassette exchange and confers resistance to Geneticin (G418).

Initial transient expression screenings revealed, that the EF1-α promoter used for stable expression of the protein in the CHO tagging vector is less active in Sf21 (data not shown). Therefore, it was exchanged by the promoter region of pIEX/Bac5, which contains a baculoviral immediate early promoter 1 and the hr5 enhancer as well as a downstream very late p10 promoter. The fragment was generated by PCR-primers flanked by BglII and HindIII and integrated in the pEF-FS-EGFP-dneo backbone where the corresponding region was deleted. A schematic map of the resulting vector termed pIE10-eGFP-dneo is shown in Figure 3-42. A second tagging plasmid was additionally constructed by exchanging the eGFP reporter gene against tdTomato. Its higher quantum yield compared to eGFP should help to identify stable producer cell lines.

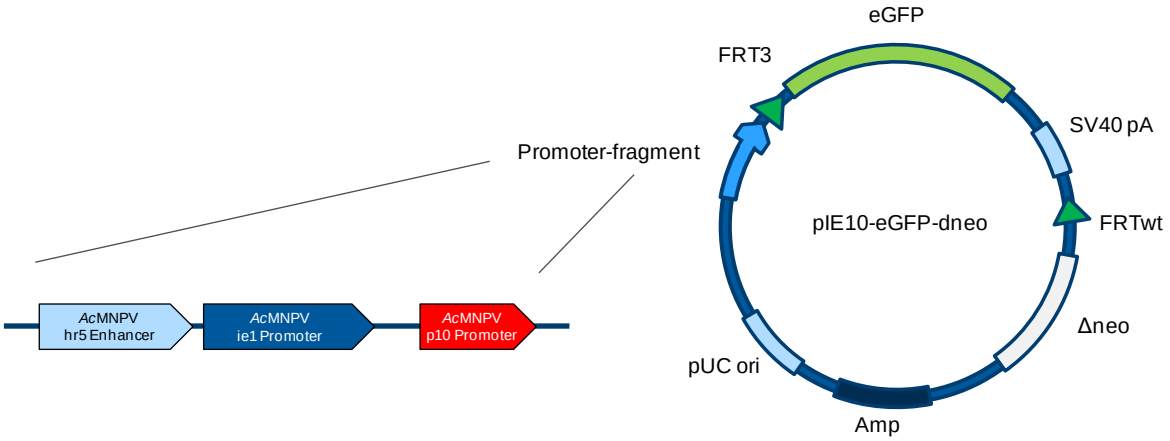


Figure 3-42: Tagging vector for the generation of stable insect cell lines
The promoter region consists of the baculoviral *Ac*MNPV immediate early promoter ie1 and its enhancer hr5 as well as the very late p10 promoter, which is only active under the presence of baculoviral transcription factors and mRNA polymerase. The marker gene eGFP is flanked by heterospecific FRT sites (FRT_3 & FRTwt). A downstream non-functional aminoglycoside phosphotransferase gene (ΔneoR) confers resistance only upon the integration of a promoter and an ATG codon via cassette exchange. The vector backbone contains a bacterial origin of replication (pUC ori) and a beta-lactamase gene for selection on Ampicillin (Amp).

3.6.2 Integration of a RMCE-*locus* into Sf21

While transfection of CHO cell lines is usually performed using the Nucleofector™ (amaxa), this method has not been established for lepidopteran cell lines. Although the manufacturer offers resuspension medium and pre-programmed electroporation settings for Sf21, the cells did not survive the transfection procedure in the amaxa device. Thus, small scale chemical transfections of suspension cultures in 6-well plates had to be tested and optimised in this work. In addition to the tests of different promoters, fluorescene rates were also compared between linearised and circular expression plasmids. The integration of linearised exogenic DNA is more likely compared to condensed plasmids, whereas the cytoplasmatic half-life and thus the expression rate in the early days post transfection is significantly higher when transferring intact plasmids. An exemplary fluorescence kinetic of these extensive evaluations is shown in Figure 3-43. Typically, 4 times higher fluorescence rates were observed upon transfection with circular plasmids. However, continuous passaging resulted in a loss of fluorescence rates and an approximation between the cultures transfected with linear or circular tagging plasmids. The transfection of the host cell line with the tagging plasmids created an inhomogeneous population of transient expression and randomly tagged cells. Accordingly, the first sorting was performed 7 days post transfection to ensure a favourable ratio of the probability for stable integration and a high absolute amount of fluorescent cells.

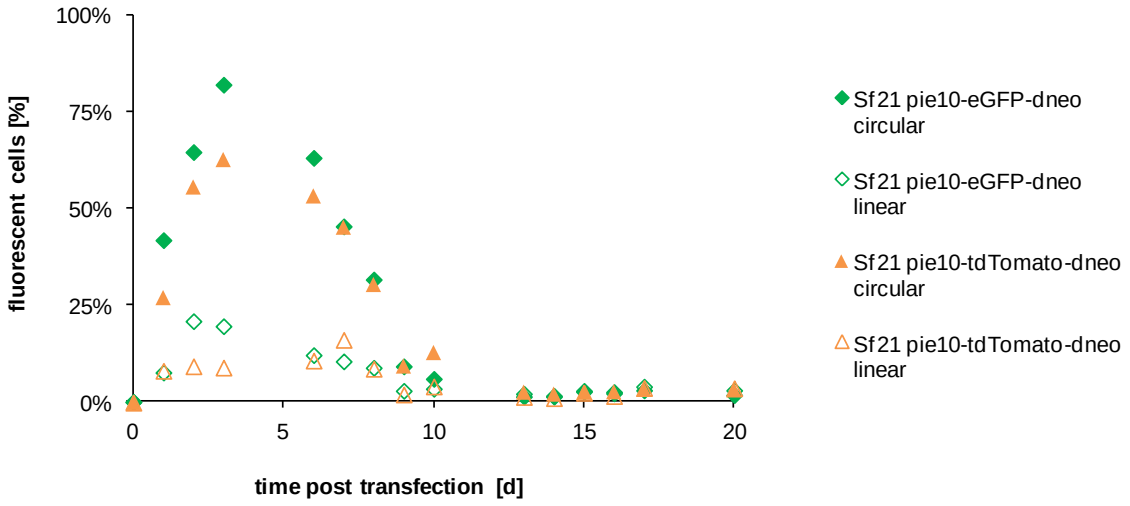


Figure 3-43: Fluorescence profile of S21 transfected with circular and linearised tagging plasmids
Suspension cultures of Sf21 were chemically transfected with linearised and circular pie10-eGFP-dneo or pie10-tdTomato-dneo respectively. Fluorescence was daily assessed at the Guava EasyCyte Mini flow cytometer at green and yellow detection channel.

3.6.3 Isolation of a clonal Sf21 cell line

The enrichment of cells that integrated the RMCE-cassette into favourable chromosomal loci was performed by at least two rounds of preparative FACS followed by a clonal isolation of cell lines from the resulting cell pool. This strategy as shown in Figure 3-44 has also been used for the generation of the CHO Lec3.2.8.1 master cell lines for RMCE in our group. By avoiding antibiotic selection a more homogenous and stable expression of the transgenes in the resulting expression cell line is achieved (Kaufman *et al.*, 2008).

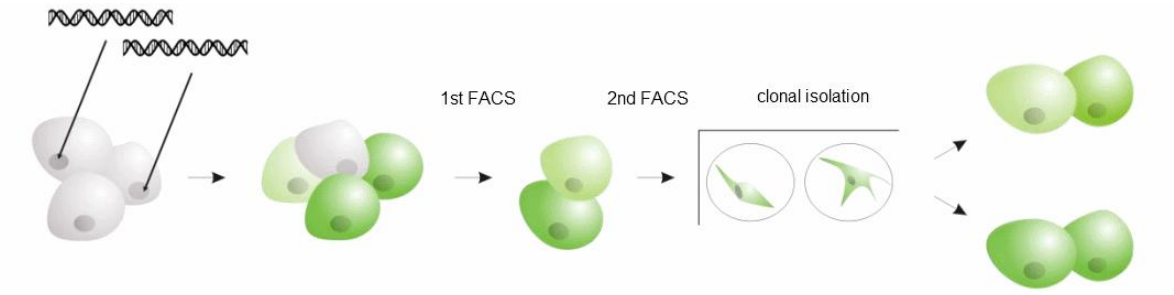


Figure 3-44: Strategy for the generation of clonal RMCE master cell lines by FACS
Random integration of the tagging vector generates an inhomogeneous pool of cells which are sorted 7 days post transfection. After expanding the cells to suspension, a second (and potentially third) sorting is performed to isolate cells out of the enriched population. Those cells are further isolated and propagated to obtain clonal master cell lines. (Modified work, original source: Wilke, 2011)

However, preparative sorting of insect cell cultures is a rare technique and not well described in the literature. In this work, FACS of insect cells transfected with the tagging vector recurrently caused a drastic decrease of the viability from 80 - 90 % before sorting to 2 - 15 % in the enriched cell pool. Although the heavily stressed cells were seeded and incubated as adherent cultures, their low viability impeded a recovery of the sorted population. As a result, it was not possible to isolate a master cell line in the initial attempts.

To reduce the shear stress during the sorting process and to enhance the viability of the sorted cell pool, supplementation with FCS and an increase in cell density up to 1×10^8 c/mL of sorted cells were tested. This represents the limit at which a proper distribution of single cells into droplets is possible. Additionally, an alternative sorter was used and a pH- and osmolarity-optimised sheath fluid was developed and evaluated (data not shown). The most promising variations to increase the viability of sorted cells were the supplementation of FCS and the use of the MoFlo high-speed cell sorter (Beckman Coulter). A viability of up to 50 % could be retrieved when cells supplemented with 10 % FCS were sorted in this device (Figure 3-45)

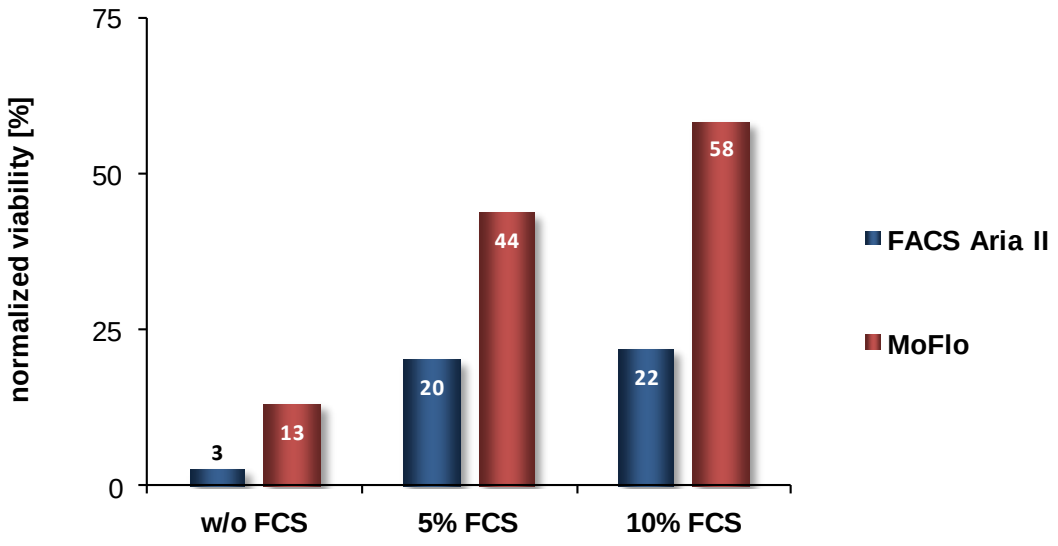


Figure 3-45: Effect of the sorting device and the supplementation with FCS on the viability of sorted cells
Sf21 were sorted 7 days post transfection with the pie10-eGFP-dneo in the BD FACS AriaII (blue) and the Beckman MoFlo sorter (red). Cultures without FCS supplements and those adapted to 5 and 10 % FCS prior to the transfection were compared. The viability measured in the re-analysis upon sorting was normalised to its value prior sorting.

Using the optimised conditions, GFP-expressing cells could be enriched from 8 % on day 7 post transfection to 94 %. The population was propagated as adherent culture for 14 days and transferred into suspension on day 21 post transfection. Strikingly, flow cytometric measurements revealed, that only ~ 20 % of the cells of this culture fluoresced (data not shown). Obviously, the integration in a significant share of the cells was either not stable or the fluorescence was still mediated by residues of transient expression. To obtain stable clonal cell lines the population was sorted again after 10 days in suspension using an automated cell deposition unit for sorting into 96-well plates. Thereby, 9 plates with single cells per well and an additional plate with 10 cells per well were inoculated. Besides the isolation of these 960 cell clones, the rest of the culture was again sorted as a population and plated in 6-well plates.

None of the cell clones in the 96-well plates recovered and formed colonies, so that the pool of sorted cells seeded at 6-well plates was consequently transferred to suspension. Again, despite the enrichment of fluorescent cells, the culture exhibited a fluorescence of only ~ 50 %. During the following 9 passages in suspension, eGFP-expression cells further decreased to ~ 30 % until they were sorted a third time 67 days post transfection. Notably, the viability of the sorted cells dramatically dropped to ~ 3 % (Figure 3-46), so that the cells had to be seeded with a very low initial density as adherent culture following the sorting. As a result, instead of growing to a confluent monolayer, small colonies were formed as shown in Figure 3-47. 11 days after the sorting, 24 of these colonies were isolated and further propagated in single wells of a 24-well plate for 18 days and subsequently transferred to suspension on day 96 post transfection of the parental Sf21 cell line. However, only one out of 20 transferred single cell clones survived the first 5 passages. Moreover, it remarkably exhibited an average fluorescence of only 30 % despite its clonal origin. Thus, further analyses of this cell line were performed as presented in chapter 3.6.5.

Subsequent to the isolation of a Sf21-derived RMCE master cell line, the generation of corresponding tagged Hi5 cell lines was additionally carried out under the author's direction as part of the Bachelor thesis of Gundula Sprick (chapter 3.6.4). In total, more than 10 iterative transfections with subsequent sorting- and isolation procedures were thus performed within this work between May 2010 and May 2011.

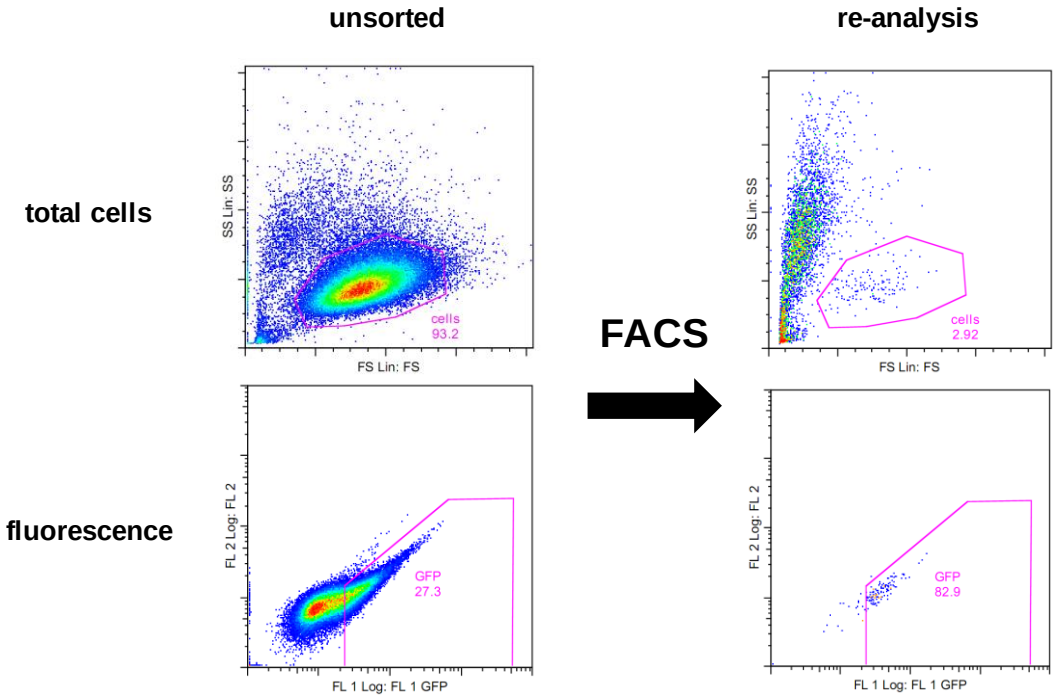


Figure 3-46: Isolation of Sf21 stably producing the eGFP reporter gene 67 days post transfection
3rd sorting of Sf21-pie10-eGFP-dneo 67 days post transfection at the MoFlo high-speed cell sorter (Beckman Coulter) in the department of Experimental immunology by Lothar Gröbe. Appropriate gates were adjusted according to negative control measurements. The pool of captured eGFP-positive cells was analysed directly after the sorting process (SS = sideward scatter, FS = forward scatter)

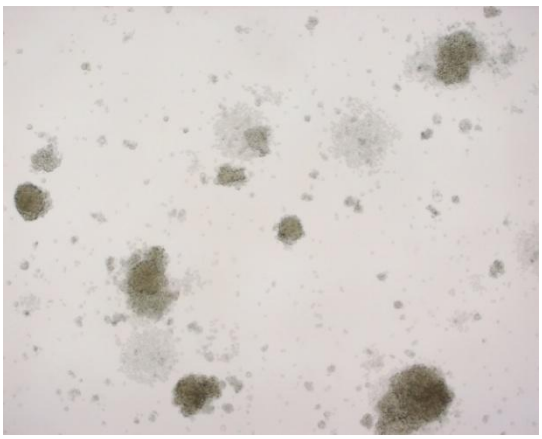


Figure 3-47: Single cell colonies of Sf21-pie10-eGFP-dneo 11 after the third sorting
Due to the low viability of cells enriched in the 3rd sorting 67 days post transfection, seeding as adherent culture in low density caused the development of colonies instead of a confluent monolayer. 20 large colonies were separated and seeded in single wells of a 24-well plate for further proliferation. Only 1 survived the first 5 passages subsequent to the transfer in suspension.

3.6.4 Generation of clonal Hi5 cell lines

For the generation of Hi5 RMCE master cell lines a second tagging vector named pie10-tdTomato-dneo harbouring tdTomato as reporter gene was used besides pie10-eGFP-dneo. The alternative reporter was used to achieve a higher fluorescence also in those cells, were the integration occurred in a transcriptionally less active locus and thereby increasing the number of fluorescent positive cells to be sorted. Figure 3-48 shows a representative enrichment of transfected Hi5 in the sorting 7 days post transfection. The resulting populations exhibited a fluorescence of more than 90 % with viabilities of ~ 25 - 30 % directly upon sorting before seeding to 6-well plates.

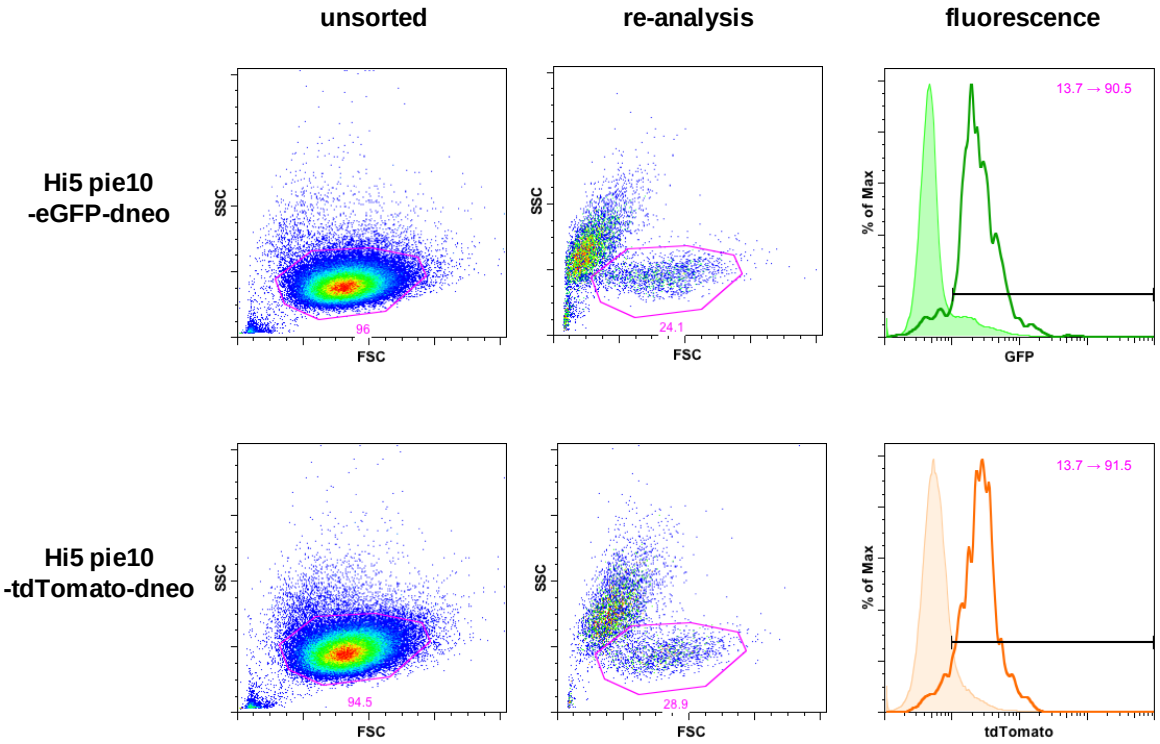


Figure 3-48: Enrichment of Hi5 cells transfected with eGFP- and tdTomato-tagging plasmids
Hi5 transfected with pie10-eGFP-dneo (upper row) and pie10-tdTomato-dneo (lower row) were sorted 7 days post transfection at the MoFlo high-speed cell sorter (Beckman Coulter) in the department of Experimental Immunology by Lothar Gröbe. Appropriate gates were adjusted according to negative control measurements. The pools of captured eGFP- and tdTomato-positive cells were analysed directly after the sorting process. Fluorescence of the culture of the culture before sorting is shown as filled areas or as surrounded light areas in the re-analysis of the sorted cultures. (SS = sideward scatter, FS = forward scatter, pink numbers indicate the percentage of the gated populations).

Following the adherent propagation for 5 days the population was transferred to suspension. Fluorescence rates of more than 50 % for cells tagged with pie10-eGFP-dneo and more than 70 % upon integration of pie10-tdTomato-dneo could be confirmed by flow cytometry. To assess the stability of the reporter gene expression in the heterogeneous cell pools, the fluorescence was analysed during continuous cultivation over 18 (Hi5-pie10-tdTomato-dneo) and 28 passages (Hi5-pie10-eGFP-dneo), respectively. Thereby, fluorescence rates of ~ 60 % could be observed 70 days post transfection in Hi5-pie10-tdToamato-dneo and 35 % 100 days post transfection in Hi5-pie10-eGFP-dneo. However, the values fluctuated between different passages (data not shown).

Since repeated preparative FACS proved to be a disadvantage, the clonal isolation of Hi5 RMCE cell lines was performed by conventional manual separation upon seeding of a dilution of the enriched population on 60 mm dishes. Colonies were picked and sequentially expanded as adherent cultures from 96-well plates up to 2 × 60 mm dishes until the cell clones could be transferred to suspension. Two tdTomato-expressing master cell lines and one with eGFP as a reporter could be generated by this means. As with the Sf21 master cell line, no homogeneous fluorescence could be detected in the Hi5 cells, although they are derived from single cell clones. Average fluorescence rate within 2 weeks of continuous cultivation in suspension were ~ 50 %for Hi5-pie10-eGFP-dneo and 58 % and 68 % for the two Hi5-pie10-tdTomato-dneo clones (data not shown).

3.6.5 Evaluation of genomic expression

Since the reporter genes within the RMCE cassettes did not have tags for purification their genomic expression in the isolated cell lines had to be analysed by fluorimetric measurements. Therefore, eGFP-RMCE master cells were infected by a virus containing mCherry to analyse the influence of the baculoviral infection on the genomic expression of eGFP within the RMCE cassette. As outlined in Figure 3-42 on page 101, the genomic expression of the reporter gene is driven by a promoter cassette comprising the baculoviral ie1 promoter and additionally by the very late p10 promoter. Upon infection of the cells with a baculovirus the latter should be activated due to viral transcription factors and the baculoviral mRNA polymerase. The progress of the fluorescence in Sf21-pie10-eGFP-dneo infected with a Bac-sumostar-mCherry virus is shown in Figure 3-49.

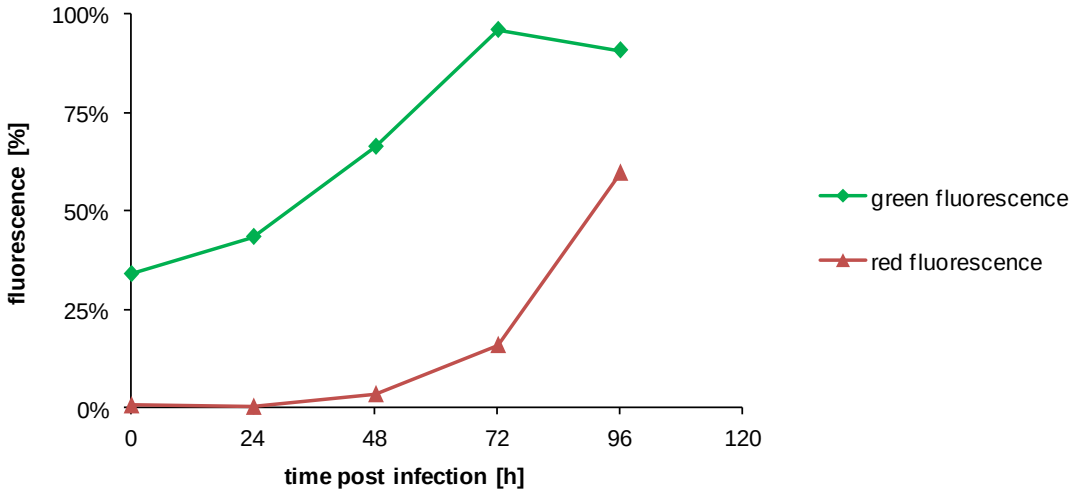


Figure 3-49: Fluorescence of Sf21-pie10-eGFP-dneo upon infection with Bac-sumo-mCherry
Sf21-pie10-eGFP-dneo was infected by a recombinant baculovirus harbouring a sumo-mCherry protein as a transfection control. Fluorescence was analysed at a Guava EasyCyte Mini flow cytometer. The fluorescence of mCherry does not affect the detection of green fluorescence since its emission wavelength is not covered by the green detector channel.

Accumulation of viral encoded mCherry is observed not before 48 h post infection which corresponds to the time window necessary for the viral DNA to enter the nucleus and start the expression genes controlled by the very late promoters. In contrast, the amount of cells which exhibit green fluorescence increases immediately upon infection and reaches a maximum of > 95 % 72 hpi. This effect is mediated by the activation of the ie1 promoter and its enhancer by immediate early viral transcription factors in the first 48 h post infection and the further increase of the transcription rate upon activation of the p10 promoter in the later phases. The fact that eGFP is expressed throughout the whole culture strikingly demonstrates that the RMCE cassette is homogenously integrated in the cells despite the low fluorescence rate observed in the cell lines in the absence of infection. A quantification of the eGFP production was performed by analysing cell lysates at a Tecan NanoQuant InfiniteM200 plate reader using an eGFP-standard for correlation. The accumulation of eGFP by constitutive genomic expression in the stable master cell lines was compared with baculoviral eGFP-expression in the non-tagged parental cell line and the enhanced expression in the master cell lines upon infection with the control virus (Figure 3-50).

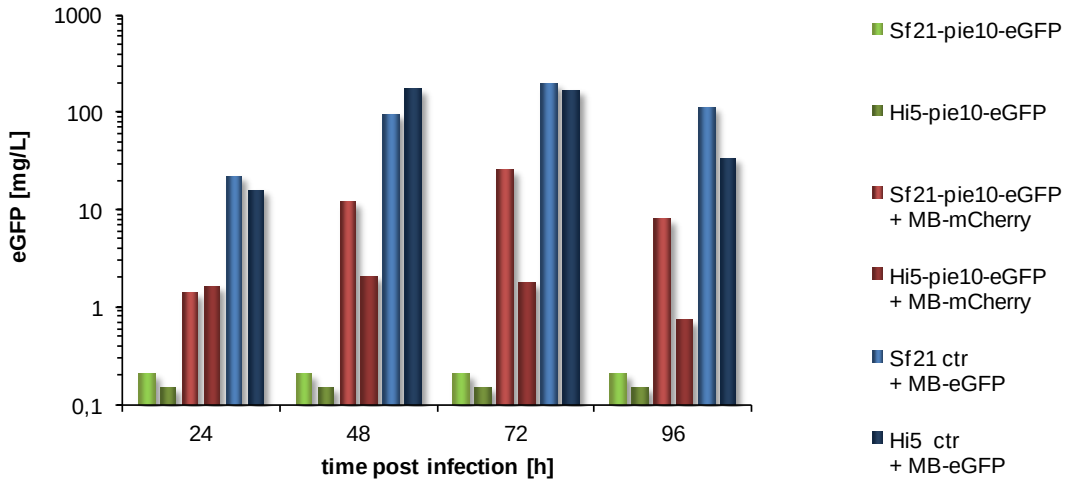


Figure 3-50: eGFP concentrations achieved by stable, viral and enhanced genomic expression
The constitutive expression of eGFP in RMCE master cells lines (green) was compared with the enhanced expression in the same cell lines upon baculoviral infection and subsequent activation of the very late p10 promoter (red) and entirely baculoviral expression in non-tagged cells (blue). The concentrations within cell lysates were determined using a Tecan NanoQuant InfiniteM200 plate reader and correlated with an eGFP-standard (instrument parameters used for the measurements are listed in chapter 2.7.2.).

The quantification revealed rather low constitutive expression yielding in 0.21 mg/L in Sf21-pie10-eGFP-dneo and 0.15 mg/L in Hi5-pie10-eGFP-dneo, respectively. However, upon infection with a baculovirus and the resulting transcriptional gain 2 mg/L and 25 mg/L could be obtained in Hi5-pie10-eGFP-dneo and Sf21-pie10-eGFP-dneo which corresponds to a 13 - 125 fold increase. In contrast, via direct baculoviral-derived expression of eGFP yields of 177 mg/L in Hi5 and 196 mg/L in Sf21 were achieved. However, taken into account, that the stable expression is driven by only a single (or few) genomic integrations of the eGFP gene, whereas up to several hundred copies of the recombinant viral genome are present in the cell in the late phase of the infection, the yields achieved with the tagged cell lines are remarkable. In fact, the baculovirus induced increase in stable genomic expression of eGFP yields comparable amounts of protein as obtained upon direct baculoviral expression and purification of the closely related mCherry (cf. chapter 3.5.1.). Consequently, the Sf21 and Hi5 master cell lines for RMCE generated in this work are well suited for stable expression of heterologous proteins.

4 Discussion

Minimizing the time required for cloning and screening of target constructs is a worthwhile goal for the optimisation of expression pipelines. Moreover, the use of engineered dedicated production cell lines facilitates the expression of challenging target proteins and leads to improved product yields. The aim of this work was thus the construction and evaluation of a novel expression vector which enables the flexible and fast screening for expressible target protein variants and the identification of the optimal expression system for any given target protein. By the incorporation of genetic elements for transient and viral expression as well as recombinase based targeted integration of transgenes into dedicated producer cell lines, the vectors should combine all expression systems available at the Recombinant Protein Expression group at the Helmholtz-Centre for Infection Research. Likewise, optimised lepidopteran cell lines that implement the RMCE system for inducible genomic expression of transgenes in addition to the BEVS were developed, thereby creating an integrated protein expression platform for the PSPF.

4.1 Evaluation of pFlpBtM-I

The first version of the vector comprised all elements required for the use as a donor vector for RMCE in the CHO Lec3.2.8.1 master cell line and in Tn7-based generation of recombinant bacmids of the Bac-to-Bac or MultiBac systems. Additionally, it was designed to be used for transient expression in insect cells, which is driven by the baculoviral immediate-early 1 promoter in combination with the homologous region 5 transactivator. This combination utilises the endogenous insect-cell transcriptional machinery and was originally used for early expression of recombinant proteins in baculoviral vectors (Jarvis *et al.*, 1996). The promoter region taken from the commercial pIEx plasmid provided by Novagen is complemented by an additional p10 promoter, that directs expression in the late/very late phases of baculoviral expression upon the generation of recombinant bacmids (Loomis *et al.*, 2007). The proof of concept studies revealed, that the vector is suitable for RMCE based generation of stable cell lines without any constrains. The CHOLec3.2.8.1 producer cell line isolated upon cassette exchange with pFlpBtM-I as a donor vector homogenously expressed the reporter protein and was used for comparative expression studies that will be discussed below in chapter 4.2.1.

Transient plasmid based expression of the fluorescent model proteins in insect cells could also be detected. However, only very faint signals were observed. This was especially the case for mCherry, which has an intrinsically lower brightness compared to eGFP. Although transient expression of heat shock proteins, kinases and phosholipases using the parental pIEx vector have been reported with yields between 3 µg/mL (kinases) and 48 µg/mL (HSP) (Loomis *et al.*, 2005), the weak fluorescence and the absence of prominent Western blot signals of the model proteins used in this work indicated that this promoter combination provides only low expression in the pFlpBtM backbone. Moreover, the unfavourable cost-benefit-relation of lipofection reagents limits this application to small scale. Transfection of Sf21 using the cost-efficient chemical polyethylenimine was evaluated in preceding experiments (data not shown). Although fluorescence could be detected, the transfection rates were significantly lower compared to commercial lipofection reagents for insect cells which corresponds to even lower protein yields. Therefore, transient expression with pFlpBtM-I based on the *Ac*MNPV ie1 promoter has to be considered as insufficient for preparative expression of target proteins. Consequently, new stronger endogenous promoters have to be identified and evaluated for this purpose. As this would be beyond the scope of this thesis, no further evaluations of this method were pursued.

In contrary to the plasmid based expression in insect cells a remarkably brighter fluorescence was observed when eGFP and mCherry were expressed in the BEVS upon the generation of recombinant viruses using pFlpBtM-I as a donor vector. Likewise, the overexpression of the fluorescent model proteins was clearly proven in SDS-PAGE and Western blots. Traces of mCherry detected in the insoluble fraction probably result from protein misfolding. The very high expression rate mediated by the late p10 promoter may exceed the processing capacity of the host cells. In contrary, the occurrence of mCherry in the culture supernatant is a result of virus mediated cell lysis in the late phase of the infection. The viral expression of the challenging ECDmTLR2 could also be detected by Western blots. A notable amount of the protein accumulated intracellularly upon overexpression. However, this characteristic was also observed to the same extend when the ECDmTLR2 construct had been expressed using a conventional BEVS donor vector. This has been performed in the PSPF before and was reproduced for this work in a comparative expression study with pFastBac-ECDmTLR2 and pFlpBtM-I-ECDmTLR2 (data not shown). Expression of TLR domains usually is problematic and yields in

improper folding so that soluble expression is only achieved as fusion proteins (Jin & Lee, 2008). Besides unfavourable construct design, an impaired secretory pathway in the late phase of baculoviral infection further contributes to this effect. The shut down of the host cell protein synthesis following baculovirus infection may lead to a limitation in the supply of secretory assistance factors (Jarvis & Garcia, 1994). Likewise, insufficient levels of ER chaperones are a limiting factor in producing functional heterologous secreted proteins in the baculovirus system (Hsu & Betenbaugh, 1997; Ailor & Betenbaugh, 1999). Both, infection kinetics and expression profiles of the performed experiments indicate, that the vector architecture of pFlpBtM-I has no negative effect on the virus replication and expression. Likewise its applicability to serve as a donor for RMCE, it is thus very well suited to be used as a BEVS donor plasmid. However, since efficient capability for fast transient construct screening was one important requirement for the novel vector besides its use in stable and baculoviral expression, a remodelling leading to the generation of pFlpBtM-II was necessary.

4.2 Evaluation of pFlpBtM-II

HEK293 EBNA cells have been successfully utilised as expression hosts for protein production for crystallography (Nettleship *et al.*, 2010). Since they are available at the Helmholtz PSPF, the improved version of pFlpBtM consequently should implement this expression system for transient expression. To utilise the EBV oriP mediated episomal replication techniques in HEK293-6E, which allow for significantly higher protein yields in transient expression, a new vector backbone comprising this replication origin had to be integrated. Likewise, the incorporation of a CMV promoter was essential to drive protein expression in mammalian cells. This was achieved via an exchange of the promoter region by a triple promoter construct derived from the vector pTriEx. It comprises a CMV promoter/enhancer combination, the same *Ac*MNPV p10 promoter as in the previous version of pFlpBtM and an additional T7 promoter controlled by a lac operator.

Proof of concept studies with the revised version in the BEVS, RMCE in CHO Lec3.2.8.1 and HEK293-6E were successful. High transfection rates of HEK293-6E and strong overexpression of the fluorescent model proteins demonstrate that the transient

expression in HEK293-6E is far superior compared to ie1/hr5 mediated expression in lepidopteran cells. Fluorescence intensities and protein signals in SDS-PAGE reached similar levels as in baculoviral expression. In stable or baculoviral expression no differences could be observed compared to pFlpBtM-I. This fact demonstrates that the modifications in the backbone neither impair the RMCE in the CHO master cell line nor influence the Tn7 transposition in the generation of recombinant bacmids in *E. coli*.

While the promoter region of pFlpBtM-I consists of only baculoviral sequences, the triple promoter integrated in pFlpBtM-II also contains the immediate early 1 promoter from the mammalian CMV and the bacteriophage T7 promoter. This has no effect on the genomic expression in stable master cell lines upon RMCE, since the FRT site is downstream of the promoter region which is thus not integrated into the host cell genome. Likewise, the tripartite promoter has no positive or negative impact on the expression in the BEVS as the CMV ie1 promoter is not active in lepidopteran cells (Pfeifer *et al.*, 1997; He *et al.*, 2008) and with a length of only 17 bp the bacteriophage T7 promoter is too small to act as a transcriptional element in eukaryotes. The baculoviral ie1 promoter, which contributes to the expression in pFlpBtM-I derived bacmids in the early stage of the infection, is absent in pFlpBtM-II. However, since its activity is negligible compared to the very late P10 its absence has no detectable influence on the protein yield. The performance of the triple promoter fragment was also tested at the Oxford Protein Production Facility (OPPF). Although no published data is available on comparative expression characteristics, no indication was found that the combination of the different promoters has any influence on their transcriptional activity compared to the use of each promoter in an isolated context (Nettleship, personal communication).

4.2.1 Expression and purification of model proteins

Subsequent to the generation and the successful proof of concept of pFlpBtM-II, its performance as a standard eukaryotic expression vector for the PSPF was demonstrated by the production of different model protein classes using pFlpBtM-II as a donor or expression vector. The fluorescent reporter mCherry, a secretory scFv-hFc fusion construct and the challenging target protein ECDmTLR2 were expressed in HEK293-6E, CHO Lec3.2.8.1 and the BEVS which enabled the identification of an optimal production strategy for each model by quantitative analyses of the yield and quality of purified proteins.

Intracellular accumulated mCherry was successfully purified by Ni-NTA IMAC upon chemical cell lysis. Average yields of 52 mg/L in transient expression and 42 mg/L in the BEVS show, that both systems are equally eligible for the production of this model protein. Due to the slightly higher yields and the significantly lower practical effort, the plasmid based transient expression in HEK293-6E is of advantage compared to the laborious BEVS in this case. However, if repeated experiments or scalability is desired, the BEVS has to be considered as superior, since viral stocks can be generated for repeated experiments with low batch-to-batch variability. Furthermore upscaling is possible without the requirement of large amounts of plasmid DNA. The single-copy integration of the transgene in stable genomic production in CHO Lec3.2.8.1 yielded lower amounts of only about 8 mg/L compared to expression with higher copy number in viral and plasmid-based expression. This corresponds to approximately 20 - 25 % of the yields achieved in the viral and transient strategies. This significant difference renders the stable expression as not favourable for the expression of "easy-to-express" target proteins with similar expression characteristics since time and labor expenses for the generation of producer cell lines would be unreasonably high.

To evaluate the expression of the scFv-hIgG1Fc construct, it was produced in the BEVS and in HEK293-6E in this work. Immunoglobulins have been produced in insect cells since the early 1990 when recombinant antibodies were usually expressed in stable myeloma cell lines. It was only by the end of that decade that transient plasmid based expression of IgGs in mammalian cells driven by viral promoters such as CMV was developed. However, only very low yields of 0.4 mg/L in COS7 were reported (Norderhaug *et al.*, 1997). In contrast, the first fully assembled IgG produced in the BEVS was captured at a level of 5 mg/L from supernatants of Sf9 monolayer cells (Hasemann & Capra, 1990). Various IgG antibodies have been produced with yields up to 18 mg/L in the following years (Verma *et al.*, 1998; Liang *et al.*, 2001). For alternative formats like scFv even higher titers were achieved and the development of optimised bioreactor processes allowed yields of up to 32 mg/L of secreted monomeric anti-phOx scFv (Kretzschmar *et al.*, 1996). Today, industry-scale production of antibodies is usually performed in stable CHO cell lines with yields of several g/L (Durocher & Butler, 2009). However, the advent of highly effective transient expression in HEK293 by the development of genetically engineered producer cell lines and cheap transfection methods using PEI lead to a 1000-fold increase in the productivity achieved by plasmid based transient expression in

mammalian cell lines (Backliwal *et al.*, 2008a), which has consequently regained more significance. Accordingly, major differences were observed in the expression profile between these systems as opposed to mCherry. Not more than 5 mg/L could be captured from insect cell supernatants by protein A chromatography upon viral expression in Sf21 and Hi5. Although the value is rather low, it still is perfectly in line with the reported titers usually achieved in the BEVS. In contrary, a significantly higher yield of 90 mg/L purified protein was obtained by transient expression in HEK293-6E, using pFlpBtM-II(beta). Parallel expression of the model protein was performed to benchmark the expression capability of pFlpBtM-II compared to the conventional pCMV vector and pTT5 which is the dedicated expression plasmid for this cell line. Due to its multiple genetic elements pFlpBtM-II(beta)-scFv-hIgG1Fc is 40 % larger than pTT5-scFv-hIgG1Fc and 30 % larger compared to pCMV-scFv-hIgG1Fc. Despite the resulting noteworthy decrease of the gene dose, still 85 % of the yield achieved with pTT5 is reached. Moreover, the concentration of captured antibody was 30% higher compared to pCMV, which is lacking the EBV-oriP. These findings impressively demonstrate that pFlpBtM-II is fully competitive to alternative expression plasmids for the use as expression plasmid in HEK293-6E. Its slightly lower expression capability compared to smaller, optimised expression plasmids for this system is negligible and certainly offset by its versatility. The fact that transient expression in HEK293-6E enables the production of at least 35 - 50 times more protein than the BEVS with less practical effort in a shorter time also underlines why the baculoviral expression has lost its former significance for fast and flexible production of immunoglobulins.

Major differences in the expression profiles between BEVS and HEK were also observed for ECDmTLR2. Viral production resulted in significant intracellular accumulation of the protein in the soluble and insoluble fractions. However, secreted protein was successfully captured by affinity chromatography. In contrary, production of ECDmTLR2 was less successful in HEK293-6E. Expression could only be detected in Western blots and the protein remained almost completely in the insoluble fraction. Traces of recombinant protein secreted to the supernatant could be enriched by Ni-NTA, but were again only detected by immunostaining. Thus, the yield was insufficient to perform further polishing steps. These findings indicate, that plasmid based expression is too weak to produce sufficient amounts of this protein species, although expression is mediated by the strong CMV promoter. Stable genomic expression proved to be superior to transient or viral

expression strategies for ECDmTLR2. Purification of the model protein from concentrated culture supernatants of the CHO producer cell line yielded notable amounts of the protein between 1 – 5 mg/L. Moreover, the SDS-PAGE gels of the eluate fractions indicate, that capturing of stably expressed and secreted ECDmTLR2 from cell culture supernatants provided a higher purity with less contamination compared to the expression in BEVS, where a significant contamination by host cell proteins due to cell lysis is observed. Regarding yield and quality, the lower expression rate in stable genomic expression is obviously compensated by proper folding and secretion in this system. Additionally, the high yields and the significantly higher quality can reproducibly be achieved and easily scaled up in bioreactor productions. The multiparallel expression screening enabled by pFlpBtM-II thereby allowed the generation of the most suitable expression strategy for the challenging protein ECDmTLR2. Consequently, the stable CHO Lec3.2.8.1 producer cell line generated with the multi-host expression vector is by now the best expression system for ECDmTLR2 available as it provides 4 – 20-fold higher protein yields compared to previously reported expression methods (Kang *et al.*, 2009).

4.2.2 Applicability of pFlpBtM-II in alternative expression hosts

It has not escaped the author's notice, that the revised promoter cassette also provides the incorporation of the CMV promoter into the bacmids via Tn7 transposition. This renders the resulting recombinant baculoviruses suitable for mediating protein expression in mammalian cells (Hofmann *et al.*, 1995). Although baculoviruses are capable of entering a broad range of mammalian cells they do not replicate in these hosts (Kost & Condreay, 2002). Therefore, baculoviruses harbouring mammalian promoters have been used for many years for the transduction of mammalian cells especially in drug discovery and assay development (Kost *et al.*, 2010) but also for large bioreactor scale production of recombinant proteins in HEK293 (Jardin *et al.*, 2008). However, high MOIs of 10 – 10,000 are required for efficient transduction so that a reconcentration of baculoviral stocks is essential (Andersson *et al.*, 2007). As this procedure is time consuming and laborious and baculovirus infected mammalian cells are not used for the production of crystallization targets in the PSPF, the evaluation of pFlpBtM-II-derived viruses in mammalian cells was omitted.

Although pFlpBtM-II is intended to be employed as a eukaryotic expression vector, the inducible T7/lac promoter included in the promoter fragment should also confer the applicability of pFlpBtM-II in *E. coli*. However, the decision whether a protein might be expressible in prokaryotic system or may require the eukaryotic PTM machinery is usually made prior to the cloning process. Moreover, many expression vectors for protein expression in *E. coli* which feature different promoter regulation principles are available and have been constantly optimised over the last decades (e.g. pET system by Novagen or pGEX by GE Healthcare). A 6.8 kb eukaryotic expression plasmid consequently is not the vector of choice for recombinant production of a prokaryotic target protein. That is why the test expressions of the fluorescent marker in *E. coli* using pFlpBtM-II were not presented in detail in this work. However, such evaluations have been performed with striking results. No overexpression of the reporters was achieved neither in Rosetta2 DE3 where T7 mRNA polymerase is controlled by an IPTG inducible lac promoter nor in BL21 AI, in which the production of T7 mRNA is driven by an arabinose promoter. Weak, basal expression was detected on Western blots instead. Since no pLys strains have been used in which T7 mRNA polymerase activity is suppressed and pFlpBtM does not contain the *lac*I repressor, the expression must be inhibited by a different mechanism. Control experiments using conventional pET vectors and the pTriEx performed in our group by Claudia Wylegalla proved both, the expressibility of the mCherry and eGFP genes in *E. coli* and the functionality of the triple promoter fragment.

A probable explanation for the inhibited protein synthesis in *E. coli* is the formation of mRNA secondary structures in the 5′ untranslated region (UTR). In the parental pTriEx vector from which the promoter fragment was obtained the Shine-Dalgarno sequence is directly downstream of the promoter section whereas additional 78 bp comprising restriction sites and the Flp recombinase target site are present in pFlpBtM-II. As shown in Figure 1-6 on page 18 the FRT site consists of 3 inverted repeats and is thus highly amenable for intramolecular pairing. To assess the effect of this sequence an mRNA secondary structure prediction was performed. It confirmed the formation of a compact fold with two distinctive hairpins as shown in Figure 4-1.

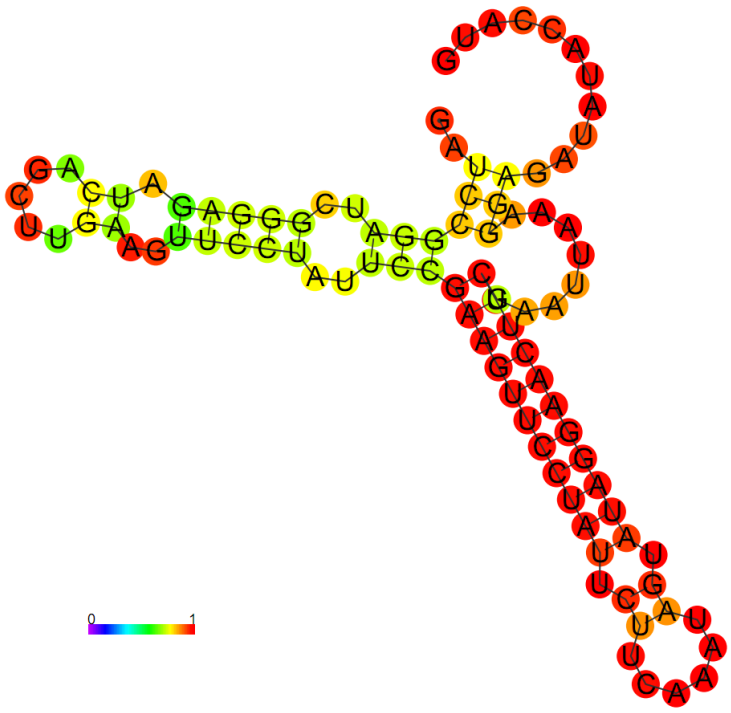


Figure 4-1: Minimal free energy mRNA secondary structure prediction of the 5'UTR of pFlpBtM-II
The minimal free energy structure of the 5'UTR of pFlpBtM-II was predicted using the RNAfold web server provided by the Institute for Theoretical Chemistry at the University of Vienna. The color code represents the base pairing probability from 0 % (purple) to 100 % (red).

The formation of this thermodynamic favourable condensed structure is driven by a notable calculated minimal free energy of ΔG = - 31.49 kcal/mol. While the eukaryotic translational machinery evidently is capable of initiating the ribosomal assembly, this strong mRNA secondary structure probably inhibits the protein production in *E. coli* effectively. In this prokaryotic host a drastic decrease of the translation initiation efficiency to 30% or less of the full expression is already observed when mRNA secondary structures are formed in the flanking region with a ΔG of ≤ - 5 kcal/mol (Kubo *et al.*, 1990). The calculated value of the predicted structure of the UTR containing the FRT site is 6 times higher, what clearly inhibits efficient expression of the target gene.

4.2.3 Screening of expressible ECDhTLR5 variants

TLR5 activates an innate immunity response upon binding of monomeric bacterial flagellin. However, the analysis of the structural basis of that recognition is impeded by the problematic expression of the receptor protein. The efforts to produce the extracellular domain of human TLR5 in the BEVS and by transient expression demonstrated that neither the full length domain nor a fusion construct containing the carboxy-terminus of soluble rainbow-trout TLR5 can be expressed in soluble form in sufficient amounts. Therefore, the new expression vector was utilised for the fast screening of different VLR hybrid proteins and truncated variants of the target protein. The hybrid constructs were generated according to the LRR hybrid technology as reported for the production of the ECD of murine TLR2 (Kim *et al.*, 2007; Jin & Lee, 2008). In total, 8 novel constructs, comprising 6 VLR hybrids and 2 truncated variants based on functional data and structure prediction were generated and evaluated for their expressibility in transient expression in HEK293-6E.

However, none of these constructs yielded significant amounts of soluble material either. The results of parallel expression tests of ECDmTLR2 demonstrated that plasmid based transient expression in HEK293-6E might be of disadvantage for this protein class. Accordingly, recombinant bacmids were generated with the most promising constructs to evaluate their expressibility in the BEVS as well. The parallel generation was facilitated by the applicability of pFlpBtM-II for both expression methods so that no recloning was necessary. The amount of soluble material obtained from the baculoviral expression of the VLR hybrids was however not enhanced compared to the initial constructs. Initial efforts to recover intracellular accumulated protein by affinity chromatography upon cell lysis remained without success (data not shown). Altogether, 12 different constructs, comprising full length ECD variants with different secretion signals, rainbow-trout- and VLR-hybrids as well as C-terminally and N-terminally truncated constructs were tested for soluble expression in this work. Due to limitations in time and workload available for this side project, the generation of further protein variants had to be omitted. Likewise, no stable CHO Lec3.2.8.1 producer cell lines were generated via RMCE as no worthwhile candidates could be identified among the tested constructs.

A very recently published paper, available only a few weeks before this thesis was submitted, reported the expression of a truncated VLR-hybrid variant of a TLR5 ortholog from the zebrafish (Yoon *et al.*, 2012). The authors strikingly emphasise that the "*key roadblock against TLR5-FliC interaction studies has been the formidable technical challenge in the expression of the mammalian TLR5 ectodomain.*" They failed in the expression of ECDTLR5 from human, mouse, frog, and trout in the BEVS and also in HEK293 which they utilised for expression tests of human and mouse ECDTLR5. Moreover, the yield of the full length zebrafish TLR5 ectodomain in baculoviral expression was insufficient for structural studies. To improve protein yield and crystallizability, the group at the Scripps Research Institute also applied the VLR hybrid method. However, sufficient expression was still only achieved with the three shortest chimeras, including TLR5 residues 22-181, 22-342 and 22-390. Notably, the largest of these fragments covers only 18 LRRs and is thus missing more than 35 % of the LRRs present in the full length ectodomain. Thereby, it is even shorter by 20 % compared to the smallest variant expressed in this work (Figure 4-2).

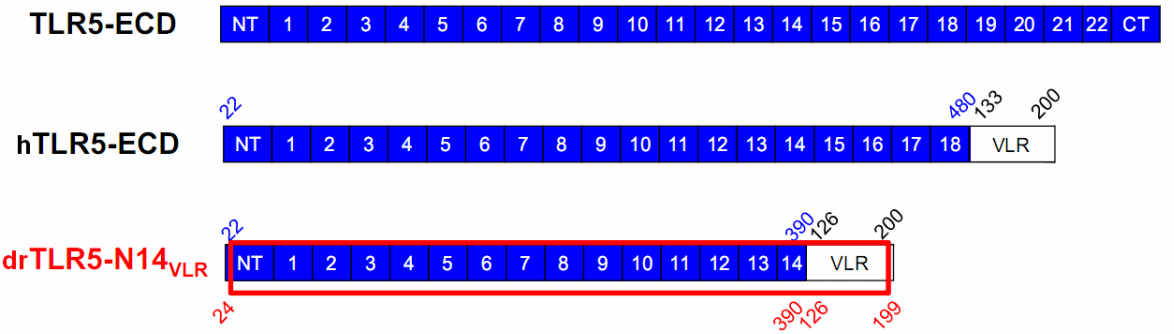
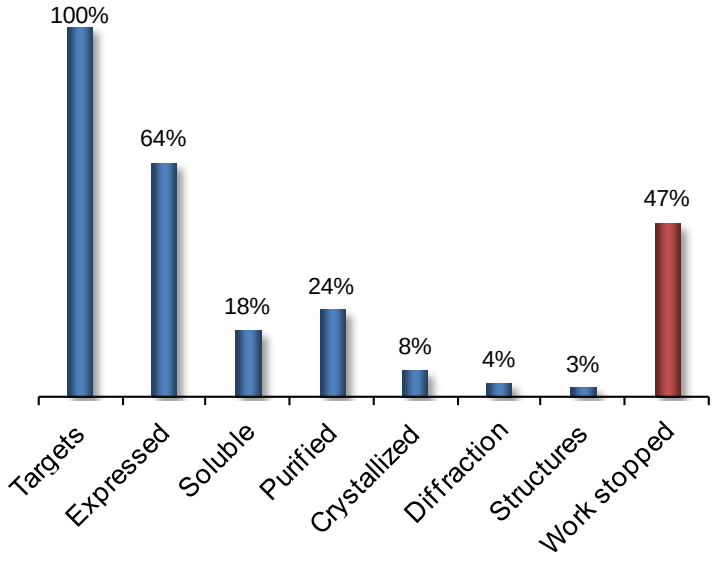


Figure 4-2: Schematic comparison of full TLR5 constructs
The full length ECD of TLR5 consists of 22 LRR motifs (blue boxes) flanked by N- and C-terminal regions (top). The shortest VLR hybrid construct of human TLR5 expressed in this work ranged to aa 480 and comprised 18 LRRs (middle). The succesfully expressed and crystallised TLR5 from the zebrafish (*Danio rerio*, drTLR5) was significantly shorter and contained only 14 LRRs (bottom). The C-terminal VLR B.61 capping region is shown as white boxes. The red frame represents the boundaries of the models built in the crystal structures. Residue numbers at N/C-terminal ends of the fragments are indicated above the module boxes (modified work based on Yoon, *et al.* 2012).

The fact that only a few constructs could be purified and crystallised is characteristic for the challenge in recombinant expression of eukaryotic proteins. Although multiple constructs of each target protein are routinely designed based on bioinformatic prediction (Graslund *et al.*, 2008) or random truncation libraries generated by automated PCR as performed at the ESPRIT platform at the EMBL in Gernoble (Yumerefendi *et al.*, 2010),

obtaining soluble material remains a major bottleneck in structural biology. Only 2/3 of all target proteins for structural investigation were found to be expressible, and only less than 20 % could be obtained with serious effort in soluble form (Figure 4-3). Accordingly, the structural genomics consortium (SGC) officially aims at a structure determination success rate of 15 - 20 % for their target proteins (Savitsky *et al.*, 2010).



Status	Number
Targets	201,446
Expressed	127,927
Soluble	36,701
Purified	47,459
Crystallized	15,117
Diffraction	7,805
Structures	5,224
Work stopped	95,127

Figure 4-3: Structural target proteins statistics
The values were obtained from the Protein Structure Initiative (PSI) TargetDB as of March, 1st 2012 and include data from structural genomics centers worldwide listed in the PSI. The value for solved structures only contains crystal structures. (Source: http://targetdb.sbkb.org/statistics/TargetStatistics.html)

The production of challenging target proteins therefore requires a multiple step strategy, including combinatorial methods to design a set of several different constructs comprising mutants, orthologs, truncations and chimeric proteins as well as a profound screening in different expression systems. The tremendous workload using such elaborate techniques and tools underlines the need for optimised expression systems such as pFlpBtM-II and the usefulness of automated high-throughput pipelines. Likewise, dedicated host cell lines, which facilitate the expression of binding partners or auxiliary proteins, represent a valuable tool to further increase the success rate.

The coexpression of chaperones might specifically aid the recombinant overexpression of Toll like receptors. Several reports characterise the glucose-regulated protein 94 (grp94) as an essential molecular chaperone for Toll-like receptors and integrins (Randow & Seed,

2001). It is a paralogue of the cytosolic HSP90 in the lumen of the endoplasmatic reticulum and is also known as HSP90b1, ErP99, gp96, endoplasmin, and TRA-1 (Wu *et al.*, 2012). Functional studies revealed, that cells without grp94 or with catalytically inactive grp94 lack functional TLRs which are retained intracellularly in the absence of this chaperone (Bloor *et al.*, 2010). Moreover, a *drosophila* ortholog has been found, which is able to chaperone multiple murine grp94 clients including integrins alpha(4), alpha(L), and beta(2) as well as TLR2 and TLR9 in grp94 deficient murine cells (Morales *et al.*, 2009). Besides the use of the VLR-hybrid technology, the coexpression of this modulator should therefore significantly enhance both protein yield and solubility of recombinantely produced TLRs in insect cells. To address this, insect master cell lines which enable the stable genomic expression of such auxiliary proteins have also been developed in this work.

4.3 Generation of lepidopteran RMCE master cell lines

Protein production in insect cells is usually performed by infecting the host cell line with recombinant baculoviruses. However, in many cases, stable genomic expression has been found to be a good alternative to the lytic baculoviral expression (Pfeifer, 1998; Kempf *et al.*, 2002; Jardin *et al.*, 2007). Moreover, BEVS host cell lines that stably express auxiliary proteins such as chaperones, kinases or binding partners can be used for a combinatorial approach for the production of demanding target protein. This principle has been applied for instance to provide BEVS host cell lines expressing human glycosyltransferases to produce glycoproteins with optimised glycan side chains (Hollister *et al.*, 2002; Hill *et al.*, 2006). However, the beneficial influence of chaperone assisted protein production in insect cells has only been demonstrated for baculoviral coexpressed chaperones yet (Ailor & Betenbaugh, 1998; Martinez-Alonso *et al.*, 2009).

In this work, engineered insect cell lines for stable genomic expression were generated by the implementation of the RMCE system to enable a fast and flexible integration of auxiliary proteins. Thereby, BEVS host cell lines that specifically improve the expression of a particular target protein can be provided by the variable genomic coexpression of appropriate modulators. Moreover, the RMCE system confers compatibility to the existing system in CHO Lec3.2.8.1 used at the Helmholtz PSPF (Wilke *et al.*, 2011).

However, the application of a FACS-based strategy for the isolation of stably tagged master cell lines proved to be difficult. In contrast to mammalian cells for which this technique has been used and optimised for many years (Nehlsen *et al.*, 2009; Zuberbuhler *et al.*, 2009), insect cells lines are obviously much more susceptible to shear stress. As a result, a severe decrease in the viability upon sorting was observed. Moreover, the tendency of isolated cells to stop proliferating compromised the propagation of single cell clones upon sorting, serial dilution and even when methylcellulose based semi-solid media were used for the isolation of colonies (data not shown). Despite this bottleneck, three RMCE master cell lines Sf21-pie10eGFP-dneo, Hi5-pie10eGFP-dneo and Hi5-pie10tdTomato-dneo were generated which stably express the corresponding marker genes. Since the screening of cell clones expressing the target gene subsequent to the cassette exchange relies solely on the G418 selection trap, the sensitivity of the lepidopteran cells for preparative sorting will not impede this technique.

Strikingly, only between 40 - 70 % of the cells of each isolated master cell line constitutively produced the markers in detectable amounts. The expression was however stable over several months. These findings suggested that in all master cell lines subpopulations exist, in which the RMCE cassette might have been deleted after the clonal isolation. However, infection assays of the master cell lines expressing eGFP as a marker revealed, that all cells of the population carry the genomic integration of the RMCE cassette. The activation of the late p10 promoter via baculoviral infection caused a significant boost in translational activity resulting in both an increased eGFP expression throughout the culture of more than 90 % and a higher fluorescence intensity. Apparently, the inhomogeneous expression might therefore result from gene silencing in a certain quota within genetically homogenous populations. The silencing of transgenes is a well-known effect, which often is mediated by cell cycle dependent histone acylation (Pikaart *et al.*, 1998; Krebs *et al.*, 1999). The integration of the RMCE cassette in different chromosomal loci with varying activity might be another explanation for the inhomogeneous expression profile of non-infected cells. This however implies that the cell lines are still non-clonal populations. In both cases, the impact of the baculoviral infection of the cell and the resulting transcriptional activation of the p10 promoter outweighs the epigenetic down- regulation of the weak constitutive expression mediated by the baculoviral ie1 promoter.

Due to the lower gene dose and the weak intrinsic promoter strength of the *Ac*MNPV ie1 promoter, the expression level of the stably integrated marker protein is 1000-fold lower than baculoviral mediated expression as determined by fluorescence quantification. However, activation of the p10 promoter mediates a 100-fold increase in the genomic expression of the integrated gene. Even though the product concentration is still 10 times lower compared the levels reached by baculoviral expression, this inducible expression profile is particularly beneficial for the production of auxiliary proteins like chaperones, kinases or phosphatases. While low levels of these proteins are provided by the constitutive ie1 activity prior to the infection, the expression level is increased when the demand for these proteins rises due to the baculoviral production of the main target. This two-step optimised expression profile is superior to conventional unbalanced coexpression, which may trigger unfavourable side effects (Martinez-Alonso *et al.*, 2010) and complicate succeeding downstream processing steps due to excessive amounts of the modulating proteins.

The inducible p10 promoter activity observed in the characterization of the stable cell lines might also be utilised for the transient screening of protein variants in unmodified insect cells. The subsequent infection of insect cells, which have been transiently transfected with either pFlpBtM-I or pFlpBtM-II provides the necessary elements to initiate p10 promoter mediated transcriptional activity of target genes encoded on the plasmid. Consequently, a boost in the expression level of these proteins is triggered. This combination of plasmid based expression driven by a strong very late baculoviral promoter should allow the production of sufficient amounts of the encoded target proteins for a construct screening without the need for the preceding generation of recombinant viruses. This technique will also allow for coexpression of different proteins by cotransfecting the host cells with several donor plasmids. By modulating the gene dose of each target in the transfection mixture, any desired stoichiometry of the target proteins can be achieved. Moreover, whenever a construct proved to be expressible in insect cells with this technique, a subsequent integration of the target gene into bacmids for reproducible production in larger scale is possible with the same pFlpBtM vector without requiring additional recloning steps.

5 Conclusions and outlook

The applicability of the novel expression and transfer vector pFlpBtM-II has been demonstrated for different production methods, including transient, viral and stable expression. A fast screening for expressible protein variants is possible in the optimised mammalian cell line HEK293-6E, but also in Sf21 and Hi5 aided by the p10 promoter activation via baculoviral superinfection. Furthermore, the plasmid can be utilised as a donor vector in all Tn7-based systems for the generation of recombinant baculoviruses for protein production in the BEVS. In combination with the CHO Lec3.2.8.1 master cell line developed at our institute and the lepidopteran master cell lines generated in this work, the plasmid also acts as a donor vector for fast and flexible RMCE based generation of cell lines. This technique can be utilised for both the production of target proteins and the deployment of auxiliary proteins for subsequent baculoviral expression.

Since the performance of the vector was demonstrated with different model protein classes it is now used as the new standard eukaryotic expression vector of the Helmholtz PSPF. The unique combination of a variety of optimised expression hosts and different production methods provides an exceptional infrastructure for protein production. This integrated platform which was termed multi-host expression system (mHost-XS) therefore is a powerful tool and will significantly fasten the timeline of expression projects at the Helmholtz PSPF. The different elements of mHost-XS and the versatile applications of pFlpBtM-II are summarised in Figure 5-1.

Several projects to provide large amounts of high quality protein for structural analyses using mHost-XS are currently in the PSPF pipeline, including further Toll like receptor variants and viral proteases. Moreover, pFlpBtM-II is in use in the labs of cooperation partners throughout Europe, e. g. the Division of Structural Biology at the University of Oxford, the Department of Biochemistry of the University of Prague, the Heinrich Pette Institute in Hamburg as well as in the structural biology research groups at the Universities of Bielefeld, Würzburg and Kiel.

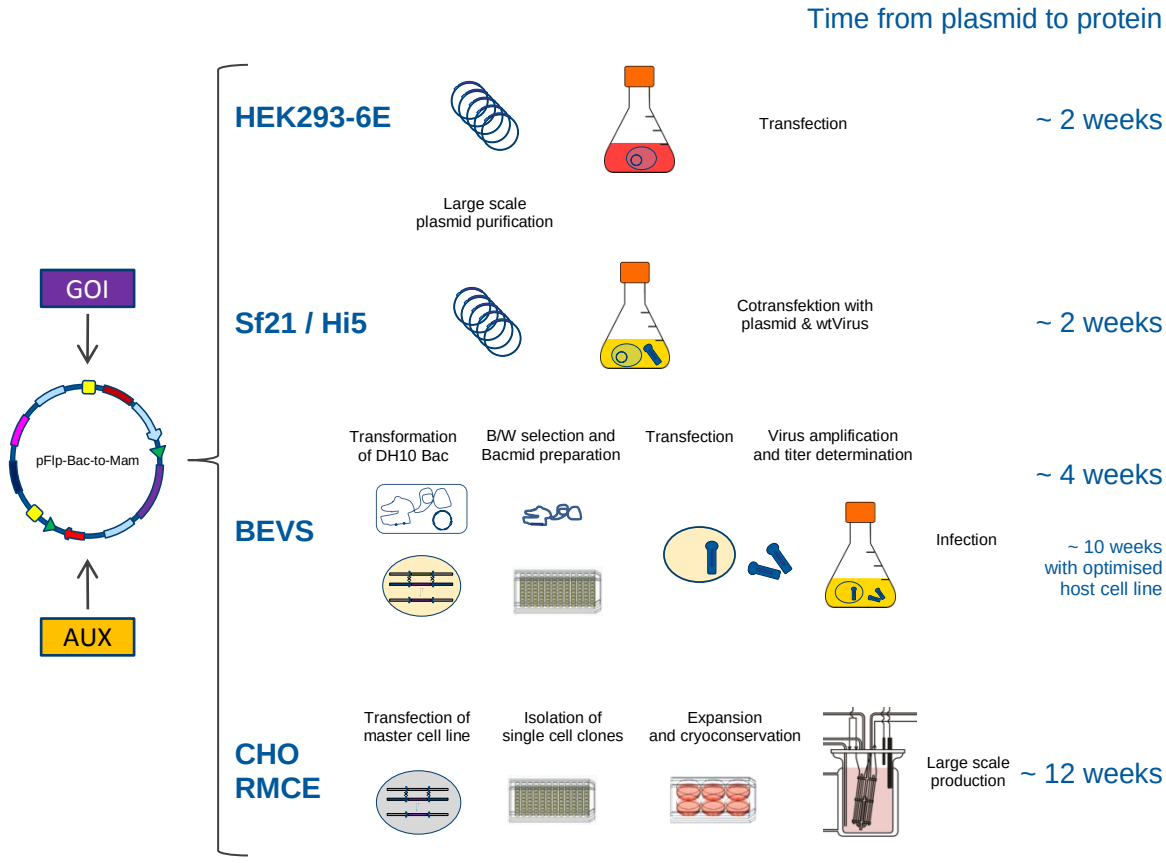


Figure 5-1: Overview of the multi-host expression system mHost-XS
Genes of interest (GOI) integrated into pFlpBtM can be produced by transient expression in HEK293-6E in which the EBV oriP provides higher product yields due to episomal replication and enhanced nuclear localization. Plasmid based transient expression is also possible in insect cell lines upon infection with a wild-type virus to initiate p10 promoter activity. The vector can also be used as a donor vector for Tn7-transposition dependent generation of recombinant bacmids for the BEVS. For stable genomic production of proteins in dedicated master cell lines the vector is capable to act as a RMCE donor vector in both CHO Lec3.2.8.1 and the novel Sf21 and Hi5 derived master cell lines. Thereby, insect cells providing genomic expression of auxiliary proteins (AUX) can be used as optimised host cells for baculoviral expression.

Projects to demonstrate the chaperone assisted expression of target proteins upon RMCE in the insect cell lines are in progress. Likewise, the technique of baculoviral enhanced transient expression upon infection of precedingly transfected cells as well as the use of recombinant baculoviruses to transduce mammalian cells will be tested at the PSPF.

The pioneering work presented here also paves the way for further optimisations of the system. The integration of multiple RMCE cassettes into host cell lines for stable coexpression of proteins is currently performed in our group. This multiplexing will require pFlpBtM variants with alternative FRT sites compatible to the new cassette. By the integration of a synthetic remodelled RMCE section into pFlpBtM-II the vector was

already prepared for such modifications by the integration of restriction sites flanking all RMCE elements like the FRT sites and promoters for their flexible exchange. It is also planned to expand mHost-XS to *Pichia pastoris* by the integration of a corresponding RMCE locus into the host cell genome. Additionally, it is worthwhile to evaluate a modification in the 5′ UTR of the vector to displace the T7 promoter downstream of the FRT site. While the short sequence is expected not to influence expression in the stable cell lines upon RMCE, the closer distance between promoter and Shine-Dalgarno sequence should enable the use of pFlpBtM for protein production in *E. coli* as well. By expanding the scope of applications of mHost-XS, the value of this tool to widen the bottleneck of producing proteins for biochemical and structural analyses will increase even further.

Bibliography

Ailor, E. & Betenbaugh, M. J. (1998) Overexpression of a cytosolic chaperone to improve solubility and secretion of a recombinant IgG protein in insect cells. *Biotechnol Bioeng,* **58,** 196-203.

Ailor, E. & Betenbaugh, M. J. (1999) Modifying secretion and post-translational processing in insect cells. *Curr Opin Biotechnol,* **10,** 142-145.

Ambinder, R. F., Mullen, M. A., Chang, Y. N., Hayward, G. S. & Hayward, S. D. (1991) Functional domains of Epstein-Barr virus nuclear antigen EBNA-1. *J Virol,* **65,** 1466-1478.

Andersen-Nissen, E., Smith, K. D., Bonneau, R., Strong, R. K. & Aderem, A. (2007) A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *J Exp Med,* **204,** 393-403.

Andersen, D. C. & Krummen, L. (2002) Recombinant protein expression for therapeutic applications. *Curr Opin Biotechnol,* **13,** 117-123.

Andersson, M., Warolen, M., Nilsson, J., Selander, M., Sterky, C., Bergdahl, K., Sorving, C., James, S. R. & Doverskog, M. (2007) Baculovirus-mediated gene transfer and recombinant protein expression do not interfere with insulin dependent phosphorylation of PKB/Akt in human SHSY-5Y and C3A cells. *BMC Cell Biol,* **8,** 6.

Andrews, B. J., Proteau, G. A., Beatty, L. G. & Sadowski, P. D. (1985) The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences. *Cell,* **40,** 795-803.

Aricescu, A. R., Assenberg, R., Bill, R. M., Busso, D., Chang, V. T., Davis, S. J., Dubrovsky, A., Gustafsson, L., Hedfalk, K., Heinemann, U., Jones, I. M., Ksiazek, D., Lang, C., Maskos, K., Messerschmidt, A., Macieira, S., Peleg, Y., Perrakis, A., Poterszman, A., Schneider, G., Sixma, T. K., Sussman, J. L., Sutton, G., Tarboureich, N., Zeev-Ben-Mordehai, T. & Jones, E. Y. (2006) Eukaryotic expression: developments for structural proteomics. *Acta Crystallogr D Biol Crystallogr,* **62,** 1114-1124.

Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr Purif,* **48,** 1-13.

Asubel, F. M., Brent, R., Kingston, R. E., Moor, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (2007) *Current Protocols in Molecular Biology.* John Wiley & Sons, New York.

Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M. & Possee, R. D. (1994) The complete DNA sequence of Autographa californica nuclear polyhedrosis virus. *Virology,* **202,** 586-605.

Backliwal, G., Hildinger, M., Chenuet, S., Wulhfard, S., De Jesus, M. & Wurm, F. M. (2008a) Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions. *Nucleic Acids Res,* **36,** e96.

Backliwal, G., Hildinger, M., Kuettel, I., Delegrange, F., Hacker, D. L. & Wurm, F. M. (2008b) Valproic acid: a viable alternative to sodium butyrate for enhancing protein expression in mammalian cell cultures. *Biotechnol Bioeng,* **101,** 182-189.

Baer, A. & Bode, J. (2001) Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr Opin Biotechnol,* **12,** 473-480.

Baldi, L., Hacker, D. L., Adam, M. & Wurm, F. M. (2007) Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnol Lett,* **29,** 677-684.

Baneyx, F. (1999) Recombinant protein expression in Escherichia coli. *Curr Opin Biotechnol,* **10,** 411-421.

Bella, J., Hindle, K. L., McEwan, P. A. & Lovell, S. C. (2008) The leucine-rich repeat structure. *Cellular and Molecular Life Sciences,* **65,** 2307-2333.

Berger, I., Bieniossek, C. & Richmond, T. J. (2008) MultiBac: Multigene Baculovirus-Based Eukaryotic Protein Complex Production. *Current Protocols in Protein Science,* **51,** 1-26.

Berger, I., Fitzgerald, D. J. & Richmond, T. J. (2004) Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol,* **22,** 1583-1587.

Berger, J., Hauber, J., Hauber, R., Geiger, R. & Cullen, B. R. (1988) Secreted placental alkaline phosphatase: a powerful new quantitative indicator of gene expression in eukaryotic cells. *Gene,* **66,** 1-10.

Bergold, G. H. (1947) Die Isolierung des Polyeder-Virus und die Natur der Polyeder. *Z. Naturfosch.,* **2,** 122-143.

Bernard, A. & Payton, M. (2001) Fermentation and growth of Escherichia coli for optimal protein production. *Curr Protoc Protein Sci,* **Chapter 5,** Unit5 3.

Berrow, N. S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N., Stuart, D. I. & Owens, R. J. (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res,* **35,** e45.

Bertani, G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *J Bacteriol,* **62,** 293-300.

Bestor, T. H. (2000) Gene silencing as a threat to the success of gene therapy. *J Clin Invest,* **105,** 409-411.

Bieniossek, C., Nie, Y., Frey, D., Olieric, N., Schaffitzel, C., Collinson, I., Romier, C., Berger, P., Richmond, T. J., Steinmetz, M. O. & Berger, I. (2009) Automated unrestricted multigene recombineering for multiprotein complex production. *Nat Methods,* **6,** 447-450.

Bishop, B., Aricescu, A. R., Harlos, K., O'Callaghan, C. A., Jones, E. Y. & Siebold, C. (2009) Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP. *Nat Struct Mol Biol,* **16,** 698-703.

Blasey, H. D. & Jäger, V. (1991) Strategies to increase the efficiency of membrane aerated and perfused animal cell bioreactors by an improved medium perfusion. *Animal cell culture and production of biologicals* (eds R. Sasaki & R. Ikura), pp. 61-73. Kluwer Academic Publishers, Dordrecht.

Blasius, A. L. & Beutler, B. (2010) Intracellular Toll-like Receptors. *Immunity,* **32,** 305-315.

Bloor, S., Maelfait, J., Krumbach, R., Beyaert, R. & Randow, F. (2010) Endoplasmic reticulum chaperone gp96 is essential for infection with vesicular stomatitis virus. *Proc Natl Acad Sci U S A,* **107,** 6970-6975.

Böedeker, B. G. D., Newcomb, R., Yuan, P., Braufman, A. & Kelsey, W. (1994) Production of recombinant factor VIII from perfusion cultures: I. Large-scale fermentation. *Animal Cell Technology: Products of Today, Prospects for Tomorrow* (eds R. W. Spier, J. B. Griffiths & W. Berthold), pp. 580-583. Butterworth-Heinemann, Oxford.

Botos, I., Segal, D. M. & Davies, D. R. (2011) The Structural Biology of Toll-like Receptors. *Structure,* **19,** 447-459.

Bouhassira, E. E., Westerman, K. & Leboulch, P. (1997) Transcriptional behavior of LCR enhancer elements integrated at the same chromosomal locus by recombinase-mediated cassette exchange. *Blood,* **90,** 3332-3344.

Boussif, O., Lezoualc'h, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B. & Behr, J. P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A,* **92,** 7297-7301.

Brake, A. J. (1990) Alpha-factor leader-directed secretion of heterologous proteins from yeast. *Methods Enzymol,* **185,** 408-421.

Branda, C. S. & Dymecki, S. M. (2004) Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev Cell,* **6,** 7-28.

Braschoss, S., Hirsch, B., Dubel, S., Stein, H. & Durkop, H. (2007) New anti-CD30 human pancreatic ribonuclease-based immunotoxin reveals strong and specific cytotoxicity in vivo. *Leuk Lymphoma,* **48,** 1179-1186.

Brondyk, W. H., Burgess, R. R. & Deutscher, M. P. (2009) Selecting an Appropriate Method for Expressing a Recombinant Protein. *Methods Enzymol*pp. 131-147. Academic Press.

Buchanan, S. G. (2002) Structural genomics: bridging functional genomics and structure-based drug design. *Curr Opin Drug Discov Devel,* **5,** 367-381.

Buchholz, F., Angrand, P.-O. & Stewart, A. F. (1998) Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotech,* **16,** 657-662.

Buchholz, F., Ringrose, L., Angrand, P. O., Rossi, F. & Stewart, A. F. (1996) Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Res,* **24,** 4256-4262.

Büssow, K., Scheich, C., Sievert, V., Harttig, U., Schultz, J., Simon, B., Bork, P., Lehrach, H. & Heinemann, U. (2005) Structural genomics of human proteins--target selection and generation of a public catalogue of expression clones. *Microb Cell Fact,* **4,** 21.

Butters, T. D., Sparks, L. M., Harlos, K., Ikemizu, S., Stuart, D. I., Jones, E. Y. & Davis, S. J. (1999) Effects of N-butyldeoxynojirimycin and the Lec3.2.8.1 mutant phenotype on N-glycan processing in Chinese hamster ovary cells: Application to glycoprotein crystallization. *PRS,* **8,** 1696-1701.

Carter, P. J. (2006) Potent antibody therapeutics by design. *Nat Rev Immunol,* **6,** 343-357.

Cereghino, G. P., Cereghino, J. L., Ilgen, C. & Cregg, J. M. (2002) Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Curr Opin Biotechnol,* **13,** 329-332.

Chen, Z., Borek, D., Padrick, S. B., Gomez, T. S., Metlagel, Z., Ismail, A. M., Umetani, J., Billadeau, D. D., Otwinowski, Z. & Rosen, M. K. (2010) Structure and control of the actin regulatory WAVE complex. *Nature,* **468,** 533-538.

Choe, J., Kelker, M. S. & Wilson, I. A. (2005) Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science,* **309,** 581-585.

Cockett, M. I., Bebbington, C. R. & Yarranton, G. T. (1990) High level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. *Biotechnology (N Y),* **8,** 662-667.

Cohen, S. N., Chang, A. C., Boyer, H. W. & Helling, R. B. (1973) Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci U S A,* **70,** 3240-3244.

Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene,* **173,** 33-38.

Coroadinha, A. S., Schucht, R., Gama-Norton, L., Wirth, D., Hauser, H. & Carrondo, M. J. (2006) The use of recombinase mediated cassette exchange in retroviral vector producer cell lines: predictability and efficiency by transgene exchange. *J Biotechnol,* **124,** 457-468.

Corsico, C. D. & Howard, B. H. (1990) Chloramphenicol acetyltransferase as a reporter in Mammalian gene transfer. *Methods Mol Biol,* **5,** 589-596.

Cote, J., Garnier, A., Massie, B. & Kamen, A. (1998) Serum-free production of recombinant proteins and adenoviral vectors by 293SF-3F6 cells. *Biotechnol Bioeng,* **59,** 567-575.

Cox, M. M. J. & Hollister, J. R. (2009) FluBlok, a next generation influenza vaccine manufactured in insect cells. *Biologicals,* **37,** 182-189.

Cregg, J., Cereghino, J., Shi, J. & Higgins, D. (2000) Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology,* **16,** 23-52.

Cullen, B. R. (2000) Utility of the secreted placental alkaline phosphatase reporter enzyme. *Methods Enzymol,* **326,** 159-164.

Davis, S. J., Puklavec, M. J., Ashford, D. A., Harlos, K., Jones, E. Y., Stuart, D. I. & Williams, A. F. (1993) Expression of soluble recombinant glycoproteins with predefined glycosylation: application to the crystallization of the T-cell glycoprotein CD2. *Protein Eng,* **6,** 229-232.

de Marco, A. (2011) Biotechnological applications of recombinant single-domain antibody fragments. *Microb Cell Fact,* **10,** 44.

Dean, D. A., Dean, B. S., Muller, S. & Smith, L. C. (1999) Sequence requirements for plasmid nuclear import. *Exp Cell Res,* **253,** 713-722.

Dean, D. A., Strong, D. D. & Zimmer, W. E. (2005) Nuclear entry of nonviral vectors. *Gene Ther,* **12,** 881-890.

Derouazi, M., Girard, P., Van Tilborgh, F., Iglesias, K., Muller, N., Bertschinger, M. & Wurm, F. M. (2004) Serum-free large-scale transient transfection of CHO cells. *Biotechnol Bioeng,* **87,** 537-545.

Dübel, S. (2007) Therapeutic Antibodies - From Past to Future. *Handbook of Therapeutic Antibodies: Technologies, Emerging Developments and Approved Therapeutics* (ed S. Dübel), pp. 1-15. WILEY-VCH Verlag.

Dulbecco, R. (1952) Production of Plaques in Monolayer Tissue Cultures by Single Particles of an Animal Virus. *Proc Natl Acad Sci U S A,* **38,** 747-752.

Durocher, Y. & Butler, M. (2009) Expression systems for therapeutic glycoprotein production. *Curr Opin Biotechnol,* **20,** 700-707.

Durocher, Y., Perret, S. & Kamen, A. (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res,* **30,** E9.

Emmons, C. & Hunsicker, L. G. (1987) Muromonab-CD3 (Orthoclone OKT3): the first monoclonal antibody approved for therapeutic use. *Iowa Med,* **77,** 78-82.

Felgner, J. H., Kumar, R., Sridhar, C. N., Wheeler, C. J., Tsai, Y. J., Border, R., Ramsey, P., Martin, M. & Felgner, P. L. (1994) Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. *J Biol Chem,* **269,** 2550-2561.

Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. & Danielsen, M. (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci U S A,* **84,** 7413-7417.

Fischer, N., Kremmer, E., Lautscham, G., Mueller-Lantzsch, N. & Grasser, F. A. (1997) Epstein-Barr virus nuclear antigen 1 forms a complex with the nuclear transporter karyopherin alpha2. *J Biol Chem,* **272,** 3999-4005.

Forneris, F., Ricklin, D., Wu, J., Tzekou, A., Wallace, R. S., Lambris, J. D. & Gros, P. (2010) Structures of C3b in complex with factors B and D give insight into complement convertase formation. *Science,* **330,** 1816-1820.

Fulwyler, M. J. (1965) Electronic separation of biological cells by volume. *Science,* **150,** 910-911.

Gallagher, S. R. (2001) One-dimensional electrophoresis using nondenaturing conditions. *Curr Protoc Cell Biol,* **Chapter 6,** Unit 6 5.

Gao, X. & Huang, L. (1995) Cationic liposome-mediated gene transfer. *Gene Ther,* **2,** 710-722.

Gardiner, G. R. & Stockdale, H. (1975) Two tissue culture media for production of lepidopteran cells and nuclear polyhedrosis viruses. *Journal of Invertebrate Pathology,* **25,** 363-370.

Geisse, S. & Fux, C. (2009) Recombinant protein production by transient gene transfer into Mammalian cells. *Methods Enzymol,* **463,** 223-238.

Gellissen, G. (2005) *Production of recombinant proteins - Novel microbial and eukaryotic expression systems.* Wiley-VCH, Weinheim.

Glick, B. R., Pasternak, J. J. & Patten, C. L. (2009) *Molecular Biotechnology: Principles and Applications of Recombinant DNA.* Asm Pr.

Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K. & Riggs, A. D. (1979) Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proc Natl Acad Sci U S A,* **76,** 106-110.

Goldstein, G., Schindler, J., Tsai, H., Cosimi, B. & Russem, P. (1985) A Randomized Clinical Trial of OKT3 Monoclonal Antibody for Acute Rejection of Cadaveric Renal Transplants. *New England Journal of Medicine,* **313,** 337-342.

Gould, S. J. & Subramani, S. (1988) Firefly luciferase as a tool in molecular and cell biology. *Anal Biochem,* **175,** 5-13.

Grace, T. D. (1967) Establishment of a line of cells from the silkworm Bombyx mori. *Nature,* **216,** 613.

Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol,* **36,** 59-74.

Graham, F. L. & van der Eb, A. J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology,* **52,** 456-467.

Granados, R. R., Derksen, A. C. G. & Dwyer, K. G. (1986) Replication of the Trichoplusia ni granulosis and nuclear polyhedrosis viruses in cell cultures. *Virology,* **152,** 472-476.

Graslund, S., Sagemark, J., Berglund, H., Dahlgren, L. G., Flores, A., Hammarstrom, M., Johansson, I., Kotenyova, T., Nilsson, M., Nordlund, P. & Weigelt, J. (2008) The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. *Protein Expr Purif,* **58,** 210-221.

Gray, D. S., Shyam (2001) Choice of Cellular Protein Expression System. *Current Protocols in Protein Science*pp. 5.16.11-15.16.34.

Grinter, N. J. (1983) A broad-host-range cloning vector transposable to various replicons. *Gene,* **21,** 133-143.

Groth, A. C., Olivares, E. C., Thyagarajan, B. & Calos, M. P. (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A,* **97,** 5995-6000.

Gurtu, V., Yan, G. & Zhang, G. (1996) IRES bicistronic expression vectors for efficient creation of stable mammalian cell lines. *Biochem Biophys Res Commun,* **229,** 295-298.

Hagemeyer, C. E., von Zur Muhlen, C., von Elverfeldt, D. & Peter, K. (2009) Single-chain antibodies as diagnostic tools and therapeutic agents. *Thromb Haemost,* **101,** 1012-1019.

Hasemann, C. A. & Capra, J. D. (1990) High-level production of a functional immunoglobulin heterodimer in a baculovirus expression system. *Proc Natl Acad Sci U S A,* **87,** 3942-3946.

He, F., Ho, Y., Yu, L. & Kwang, J. (2008) WSSV ie1 promoter is more efficient than CMV promoter to express H5 hemagglutinin from influenza virus in baculovirus as a chicken vaccine. *BMC Microbiology,* **8,** 238.

Heim, R., Prasher, D. C. & Tsien, R. Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A,* **91,** 12501-12504.

Heim, R. & Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology,* **6,** 178-182.

Hill, D. R., Aumiller, J. J., Shi, X. & Jarvis, D. L. (2006) Isolation and analysis of a baculovirus vector that supports recombinant glycoprotein sialylation by SfSWT-1 cells cultured in serum-free medium. *Biotechnol Bioeng,* **95,** 37-47.

Hoeijmakers, J. H., Odijk, H. & Westerveld, A. (1987) Differences between rodent and human cell lines in the amount of integrated DNA after transfection. *Exp Cell Res,* **169,** 111-119.

Hofmann, C., Sandig, V., Jennings, G., Rudolph, M., Schlag, P. & Strauss, M. (1995) Efficient gene transfer into human hepatocytes by baculovirus vectors. *Proc Natl Acad Sci U S A,* **92,** 10099-10103.

Holliger, P. & Hudson, P. J. (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotechnol,* **23,** 1126-1136.

Hollister, J., Grabenhorst, E., Nimtz, M., Conradt, H. & Jarvis, D. L. (2002) Engineering the protein N-glycosylation pathway in insect cells for production of biantennary, complex N-glycans. *Biochemistry,* **41,** 15093-15104.

Hosoda, F., Nishimura, S., Uchida, H. & Ohki, M. (1990) An F factor based cloning system for large DNA fragments. *Nucleic Acids Res,* **18,** 3863-3869.

Hsu, T.-A. & Betenbaugh, M. J. (1997) Coexpression of Molecular Chaperone BiP Improves Immunoglobulin solubility and IgG Secretion from Trichoplusia ni Insect Cells. *Biotechnol Prog,* **13,** 96-104.

Huh, S. H., Do, H. J., Lim, H. Y., Kim, D. K., Choi, S. J., Song, H., Kim, N. H., Park, J. K., Chang, W. K., Chung, H. M. & Kim, J. H. (2007) Optimization of 25 kDa linear polyethylenimine for efficient gene delivery. *Biologicals,* **35,** 165-171.

Hunt, I. (2005) From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expr Purif,* **40,** 1-22.

Hunte, C. & Michel, H. (2002) Crystallisation of membrane proteins mediated by antibody fragments. *Curr Opin Struct Biol,* **12,** 503-508.

Janeway, C. A., Jr. & Medzhitov, R. (2002) Innate immune recognition. *Annu Rev Immunol,* **20,** 197-216.

Jardin, B. A., Montes, J., Lanthier, S., Tran, R. & Elias, C. (2007) High cell density fed batch and perfusion processes for stable non-viral expression of secreted alkaline phosphatase (SEAP) using insect cells: comparison to a batch Sf-9-BEV system. *Biotechnol Bioeng,* **97,** 332-345.

Jardin, B. A., Zhao, Y., Selvaraj, M., Montes, J., Tran, R., Prakash, S. & Elias, C. B. (2008) Expression of SEAP (secreted alkaline phosphatase) by baculovirus mediated transduction of HEK 293 cells in a hollow fiber bioreactor system. *J Biotechnol,* **135,** 272-280.

Jarvis, D. L. & Garcia, A., Jr. (1994) Biosynthesis and processing of the Autographa californica nuclear polyhedrosis virus gp64 protein. *Virology,* **205,** 300-313.

Jarvis, D. L., Kawar, Z. S. & Hollister, J. R. (1998) Engineering N-glycosylation pathways in the baculovirus-insect cell system. *Curr Opin Biotechnol,* **9,** 528-533.

Jarvis, D. L., Weinkauf, C. & Guarino, L. A. (1996) Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expr Purif,* **8,** 191-203.

Jin, M. S., Kim, S. E., Heo, J. Y., Lee, M. E., Kim, H. M., Paik, S. G., Lee, H. & Lee, J. O. (2007) Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell,* **130,** 1071-1082.

Jin, M. S. & Lee, J. O. (2008) Application of hybrid LRR technique to protein crystallization. *BMB Rep,* **41,** 353-357.

Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature,* **321,** 522-525.

Jordan, E., Hust, M., Roth, A., Biedendieck, R., Schirrmann, T., Jahn, D. & Dubel, S. (2007) Production of recombinant antibody fragments in Bacillus megaterium. *Microb Cell Fact,* **6,** 2.

Julius, M. H., Masuda, T. & Herzenberg, L. A. (1972) Demonstration that antigen-binding cells are precursors of antibody-producing cells after purification with a fluorescence-activated cell sorter. *Proc Natl Acad Sci U S A,* **69,** 1934-1938.

Kajava, A. V. (1998) Structural diversity of leucine-rich repeat proteins. *J Mol Biol,* **277,** 519-527.

Kang, J. Y., Nan, X., Jin, M. S., Youn, S. J., Ryu, Y. H., Mah, S., Han, S. H., Lee, H., Paik, S. G. & Lee, J. O. (2009) Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity,* **31,** 873-884.

Kaufman, W. L., Kocman, I., Agrawal, V., Rahn, H. P., Besser, D. & Gossen, M. (2008) Homogeneity and persistence of transgene expression by omitting antibiotic selection in cell line isolation. *Nucleic Acids Res,* **36,** e111.

Kempf, J., Snook, L. A., Vonesch, J.-L., Dahms, T. E. S., Pattus, F. & Massotte, D. (2002) Expression of the human μ opioid receptor in a stable Sf9 cell line. *Journal of Biotechnology,* **95,** 181-187.

Kim, H. M., Park, B. S., Kim, J. I., Kim, S. E., Lee, J., Oh, S. C., Enkhbayar, P., Matsushima, N., Lee, H., Yoo, O. J. & Lee, J. O. (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell,* **130,** 906-917.

Kim, M. S. & Lee, G. M. (2008) Use of Flp-mediated cassette exchange in the development of a CHO cell line stably producing erythropoietin. *J Microbiol Biotechnol,* **18,** 1342-1351.

Kim, N. S., Byun, T. H. & Lee, G. M. (2001) Key determinants in the occurrence of clonal variation in humanized antibody expression of cho cells during dihydrofolate reductase mediated gene amplification. *Biotechnol Prog,* **17,** 69-75.

King, L. A., Mann, S. G., Lawrie, A. M. & Mulshaw, S. H. (1991) Replication of wild-type and recombinant Autographa californica nuclear polyhedrosis virus in a cell line derived from Mamestra brassicae. *Virus Research,* **19,** 93-103.

Kishida, T., Asada, H., Kubo, K., Sato, Y. T., Shin-Ya, M., Imanishi, J., Yoshikawa, K. & Mazda, O. (2008) Pleiotrophic functions of Epstein-Barr virus nuclear antigen-1 (EBNA-1) and oriP differentially contribute to the efficiency of transfection/expression of exogenous gene in mammalian cells. *J Biotechnol,* **133,** 201-207.

Kito, M., Itami, S., Fukano, Y., Yamana, K. & Shibui, T. (2002) Construction of engineered CHO strains for high-level production of recombinant proteins. *Appl Microbiol Biotechnol,* **60,** 442-448.

Kitts, P. A., Ayres, M. D. & Possee, R. D. (1990) Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucleic Acids Res,* **18,** 5667-5672.

Kitts, P. A. & Possee, R. D. (1993) A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques,* **14,** 810-817.

Köhler, G. & Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature,* **256,** 495-497.

Kolb, A. F. (2002) Genome engineering using site-specific recombinases. *Cloning and Stem Cells,* **4,** 65-80.

Kontermann, R. E. (2011) Alternative antibody formats. *Curr Opin Mol Ther,* **12,** 176-183.

Kost, T. A. & Condreay, J. P. (2002) Recombinant baculoviruses as mammalian cell gene-delivery vectors. *Trends Biotechnol,* **20,** 173-180.

Kost, T. A., Condreay, J. P. & Ames, R. S. (2010) Baculovirus gene delivery: a flexible assay development tool. *Curr Gene Ther,* **10,** 168-173.

Krammer, F. & Grabherr, R. (2010) Alternative influenza vaccines made by insect cells. *Trends Mol Med,* **16,** 313-320.

Krebs, J. E., Kuo, M. H., Allis, C. D. & Peterson, C. L. (1999) Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. *Genes Dev,* **13,** 1412-1421.

Kretzschmar, T., Aoustin, L., Zingel, O., Marangi, M., Vonach, B., Towbin, H. & Geiser, M. (1996) High-level expression in insect cells and purification of secreted monomeric single-chain Fv antibodies. *J Immunol Methods,* **195,** 93-101.

Kubo, M., Higo, Y. & Imanaka, T. (1990) Biological threshold values of procaryotic gene expression which is controlled by the DNA inverted repeat sequence and the mRNA secondary structure. *Journal of Fermentation and Bioengineering,* **69,** 305-307.

Kuzio, J., Rohel, D. Z., Curry, C. J., Krebs, A., Carstens, E. B. & Faulkner, P. (1984) Nucleotide sequence of the p10 polypeptide gene of Autographa californica nuclear polyhedrosis virus. *Virology,* **139,** 414-418.

Kyhse-Andersen, J. (1984) Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J Biochem Biophys Methods,* **10,** 203-209.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227,** 680-685.

Langle-Rouault, F., Patzel, V., Benavente, A., Taillez, M., Silvestre, N., Bompard, A., Sczakiel, G., Jacobs, E. & Rittner, K. (1998) Up to 100-fold increase of apparent gene expression in the presence of Epstein-Barr virus oriP sequences and EBNA1: implications of the nuclear import of plasmids. *J Virol,* **72,** 6181-6185.

Laroche, Y., Storme, V., De Meutter, J., Messens, J. & Lauwereys, M. (1994) High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, Pichia pastoris. *Biotechnology (N Y),* **12,** 1119-1124.

Lebkowski, J. S., Clancy, S. & Calos, M. P. (1985) Simian virus 40 replication in adenovirus-transformed human cells antagonizes gene expression. *Nature,* **317,** 169-171.

Legardinier, S., Duonor-Cerutti, M., Devauchelle, G., Combarnous, Y. & Cahoreau, C. (2005) Biological activities of recombinant equine luteinizing hormone/chorionic gonadotropin (eLH/CG) expressed in Sf9 and Mimic insect cell lines. *J Mol Endocrinol,* **34,** 47-60.

Lehmann, J., Piehl, G. W. & Schulz, R. (1987) Bubble free cell culture aeration with porous moving membranes. *Dev Biol Stand,* **66,** 227-240.

Liang, M., Dubel, S., Li, D., Queitsch, I., Li, W. & Bautz, E. K. (2001) Baculovirus expression cassette vectors for rapid production of complete human IgG from phage display selected antibody fragments. *J Immunol Methods,* **247,** 119-130.

Lindner, S. E. & Sugden, B. (2007) The plasmid replicon of Epstein-Barr virus: mechanistic insights into efficient, licensed, extrachromosomal replication in human cells. *Plasmid,* **58,** 1-12.

Liu, W., Xiong, Y. & Gossen, M. (2006) Stability and homogeneity of transgene expression in isogenic cells. *J Mol Med (Berl),* **84,** 57-64.

Loomis, K., Rockwell, H., Sternard, H. & Novy, R. (2007) Dual-Purpose Insect Cell Expression Vector for Transient Transfection and Baculovirus Generation. *Association of Biomolecular Research Facilities Meeting*. Tampa, Fl.

Loomis, K., Yaeger, K., Batenjany, M., Mehler, M., Grabski, A., Wong, S. & Novy, R. (2005) InsectDirect System: Rapid, High-level Protein Expression and Purification from Insect Cells. *Journal of Structural and Functional Genomics,* **6,** 189-194.

Luckow, V. A., Lee, S. C., Barry, G. F. & Olins, P. O. (1993) Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion

of foreign genes into a baculovirus genome propagated in Escherichia coli. *J Virol,* **67,** 4566-4579.

Martignoni, M. E. (1984) Baculovirus: An Attractive Biological Alternative. *Chemical and Biological Controls in Forestry*pp. 55-67. American Chemical Society.

Martinez-Alonso, M., Garcia-Fruitos, E., Ferrer-Miralles, N., Rinas, U. & Villaverde, A. (2010) Side effects of chaperone gene co-expression in recombinant protein production. *Microb Cell Fact,* **9,** 64.

Martinez-Alonso, M., Toledo-Rubio, V., Noad, R., Unzueta, U., Ferrer-Miralles, N., Roy, P. & Villaverde, A. (2009) Rehosting of bacterial chaperones for high-quality protein production. *Appl Environ Microbiol,* **75,** 7850-7854.

Matsushima, N., Tanaka, T., Enkhbayar, P., Mikami, T., Taga, M., Yamada, K. & Kuroki, Y. (2007) Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics,* **8,** 124.

Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L. & Lukyanov, S. A. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotechnol,* **17,** 969-973.

McBurney, M. W., Mai, T., Yang, X. & Jardine, K. (2002) Evidence for repeat-induced gene silencing in cultured Mammalian cells: inactivation of tandem repeats of transfected genes. *Exp Cell Res,* **274,** 1-8.

McLellan, J. S., Pancera, M., Carrico, C., Gorman, J., Julien, J. P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K., O'Dell, S., Patel, N., Shahzad-ul-Hussan, S., Yang, Y., Zhang, B., Zhou, T., Zhu, J., Boyington, J. C., Chuang, G. Y., Diwanji, D., Georgiev, I., Kwon, Y. D., Lee, D., Louder, M. K., Moquin, S., Schmidt, S. D., Yang, Z. Y., Bonsignori, M., Crump, J. A., Kapiga, S. H., Sam, N. E., Haynes, B. F., Burton, D. R., Koff, W. C., Walker, L. M., Phogat, S., Wyatt, R., Orwenyo, J., Wang, L. X., Arthos, J., Bewley, C. A., Mascola, J. R., Nabel, G. J., Schief, W. R., Ward, A. B., Wilson, I. A. & Kwong, P. D. (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature,* **480,** 336-343.

Medzhitov, R. (2009) Approaching the asymptote: 20 years later. *Immunity,* **30,** 766-775.

Medzhitov, R., Preston-Hurlburt, P. & Janeway Jr, C. A. (1997) A human homologue of the Drosophila toll protein signals activation of adaptive immunity. *Nature,* **388,** 394-397.

Meghrous, J., Mahmoud, W., Jacob, D., Chubet, R., Cox, M. & Kamen, A. A. (2009) Development of a simple and high-yielding fed-batch process for the production of influenza vaccines. *Vaccine,* **28,** 309-316.

Menzel, C., Schirrmann, T., Konthur, Z., Jostock, T. & Dubel, S. (2008) Human antibody RNase fusion protein targeting CD30+ lymphomas. *Blood,* **111,** 3830-3837.

Monie, T. P., Bryant, C. E. & Gay, N. J. (2009) Activating immunity: lessons from the TLRs and NLRs. *Trends Biochem Sci,* **34,** 553-561.

Morales, C., Wu, S., Yang, Y., Hao, B. & Li, Z. (2009) Drosophila glycoprotein 93 Is an ortholog of mammalian heat shock protein gp96 (grp94, HSP90b1, HSPC4) and retains disulfide bond-independent chaperone function for TLRs and integrins. *J Immunol,* **183,** 5121-5128.

Moresco, E. M., LaVine, D. & Beutler, B. (2011) Toll-like receptors. *Curr Biol,* **21,** R488-493.

Mulder, G. J. (1838) On the composition of some animal substances. *Bulletin des Sciences Physiques et Naturelles en Neerlande*.

Nakamura, S., Takasaki, H., Kobayashi, K. & Kato, A. (1993) Hyperglycosylation of hen egg white lysozyme in yeast. *J Biol Chem,* **268,** 12706-12712.

Nehlsen, K., Schucht, R., da Gama-Norton, L., Kromer, W., Baer, A., Cayli, A., Hauser, H. & Wirth, D. (2009) Recombinant protein expression by targeting pre-selected chromosomal loci. *BMC Biotechnol,* **9,** 100.

Nettleship, J. E., Assenberg, R., Diprose, J. M., Rahman-Huq, N. & Owens, R. J. (2010) Recent advances in the production of proteins in insect and mammalian cells for structural biology. *J Struct Biol,* **172,** 55-65.

Niemann, H. H., Jager, V., Butler, P. J., van den Heuvel, J., Schmidt, S., Ferraris, D., Gherardi, E. & Heinz, D. W. (2007) Structure of the human receptor tyrosine kinase met in complex with the Listeria invasion protein InlB. *Cell,* **130,** 235-246.

Norderhaug, L., Olafsen, T., Michaelsen, T. E. & Sandlie, I. (1997) Versatile vectors for transient and stable expression of recombinant antibody molecules in mammalian cells. *J Immunol Methods,* **204,** 77-87.

Novy, R., Yeager, K. & Monsma, S. (1999) pTriEx-1 Multisystem Vector for protein expression in E. Coli, mammalian, and insect cells. *Innovations,* **10,** 5.

O'Gorman, S., Fox, D. T. & Wahl, G. M. (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science,* **251,** 1351-1355.

Ohkawa, T., Volkman, L. E. & Welch, M. D. (2010) Actin-based motility drives baculovirus transit to the nucleus and cell surface. *J Cell Biol,* **190,** 187-195.

Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996) Crystal structure of the Aequorea victoria green fluorescent protein. *Science,* **273,** 1392-1395.

Ospelt, C. & Gay, S. (2010) TLRs and chronic inflammation. *Int J Biochem Cell Biol,* **42,** 495-505.

Pancer, Z. & Cooper, M. D. (2006) The evolution of adaptive immunity. *Annu Rev Immunol,* **24,** 497-518.

Park, B. S., Song, D. H., Kim, H. M., Choi, B. S., Lee, H. & Lee, J. O. (2009) The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature,* **458,** 1191-1195.

Pasare, C. & Medzhitov, R. (2005) Toll-like receptors: linking innate and adaptive immunity. *Adv Exp Med Biol,* **560,** 11-18.

Perrett, D. (2007) From 'protein' to the beginnings of clinical proteomics. *Proteomics Clin Appl,* **1,** 720-738.

Pfeifer, T. A. (1998) Expression of heterologous proteins in stable insect cell culture. *Curr Opin Biotechnol,* **9,** 518-521.

Pfeifer, T. A., Hegedus, D. D., Grigliatti, T. A. & Theilmann, D. A. (1997) Baculovirus immediate-early promoter-mediated expression of the Zeocin resistance gene for use as a dominant selectable marker in dipteran and lepidopteran insect cell lines. *Gene,* **188,** 183-190.

Pham, P. L., Perret, S., Doan, H. C., Cass, B., St-Laurent, G., Kamen, A. & Durocher, Y. (2003) Large-scale transient transfection of serum-free suspension-growing HEK293 EBNA1 cells: peptone additives improve cell growth and transfection efficiency. *Biotechnol Bioeng,* **84,** 332-342.

Pikaart, M. J., Recillas-Targa, F. & Felsenfeld, G. (1998) Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev,* **12,** 2852-2862.

Qiao, J., Oumard, A., Wegloehner, W. & Bode, J. (2009) Novel tag-and-exchange (RMCE) strategies generate master cell clones with predictable and stable transgene expression properties. *J Mol Biol,* **390,** 579-594.

Queen, C., Schneider, W. P., Selick, H. E., Payne, P. W., Landolfi, N. F., Duncan, J. F., Avdalovic, N. M., Levitt, M., Junghans, R. P. & Waldmann, T. A. (1989) A humanized antibody that binds to the interleukin 2 receptor. *Proc Natl Acad Sci U S A,* **86,** 10029-10033.

Randow, F. & Seed, B. (2001) Endoplasmic reticulum chaperone gp96 is required for innate immunity but not cell viability. *Nat Cell Biol,* **3,** 891-896.

Raymond, C. S. & Soriano, P. (2007) High-Efficiency FLP and PhiC31 Site-Specific Recombination in Mammalian Cells. *PLoS ONE,* **2,** e162.

Rohrmann, G. F. (2008) *Baculovirus Molecular Biology.* National Library of Medicine, Bethesda (MD).

Sabroe, I., Parker, L. C., Dower, S. K. & Whyte, M. K. (2008) The role of TLR activation in inflammation. *J Pathol,* **214,** 126-135.

Sambrook, J. & Russel, D. W. (2000) *Molecular Cloning - A Laboratory Manual.* Cold Spring Harbour Laboratory Press, Cold Spring Harbour.

Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A. & Gileadi, O. (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol,* **172,** 3-13.

Schagger, H. & von Jagow, G. (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem,* **199,** 223-231.

Schlake, T. & Bode, J. (1994) Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry,* **33,** 12746-12751.

Schucht, R., Coroadinha, A. S., Zanta-Boussif, M. A., Verhoeyen, E., Carrondo, M. J., Hauser, H. & Wirth, D. (2006) A new generation of retroviral producer cells: predictable and stable virus production by Flp-mediated site-specific integration of retroviral vectors. *Mol Ther,* **14,** 285-292.

Scrable, H. & Stambrook, P. J. (1997) Activation of the lac repressor in the transgenic mouse. *Genetics,* **147,** 297-304.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E. & Tsien, R. Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol,* **22,** 1567-1572.

Shimomura, O., Johnson, F. H. & Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol,* **59,** 223-239.

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc Natl Acad Sci U S A,* **89,** 8794-8797.

Sorensen, H. P. & Mortensen, K. K. (2005) Advanced genetic strategies for recombinant protein expression in Escherichia coli. *J Biotechnol,* **115,** 113-128.

Southern, P. J. & Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J Mol Appl Genet,* **1,** 327-341.

Stanley, P. (1989) Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol Cell Biol,* **9,** 377-383.

Stark, W. M., Boocock, M. R. & Sherratt, D. J. (1992) Catalysis by site-specific recombinases. *Trends in Genetics,* **8,** 432-439.

Summers, M. D. (2006) Milestones leading to the genetic engineering of baculoviruses as expression vector systems and viral pesticides. *Adv Virus Res,* **68,** 3-73.

Summers, M. D. & Smith, G. E. (1987) A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures. *Texas Agricultural Experiment Station, Bulletin No. 1555*.

Takata, Y., Kondo, S., Goda, N., Kanegae, Y. & Saito, I. (2011) Comparison of efficiency between FLPe and Cre for recombinase-mediated cassette exchange in vitro and in adenovirus vector production. *Genes to Cells,* **16,** 765-777.

Tereshko, V., Uysal, S., Koide, A., Margalef, K., Koide, S. & Kossiakoff, A. A. (2008) Toward chaperone-assisted crystallography: protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold. *Protein Sci,* **17,** 1175-1187.

Thyagarajan, B. & Calos, M. P. (2005) Site-specific integration for high-level protein production in mammalian cells. *Methods Mol Biol,* **308,** 99-106.

Thyagarajan, B., Guimaraes, M. J., Groth, A. C. & Calos, M. P. (2000) Mammalian genomes contain active recombinase recognition sites. *Gene,* **244,** 47-54.

Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S. & Calos, M. P. (2001) Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Mol Cell Biol,* **21,** 3926-3934.

Tjio, J. H. & Puck, T. T. (1958) Genetics of somatic mammalian cells. II. Chromosomal constitution of cells in tissue culture. *J Exp Med,* **108,** 259-268.

Tomiya, N., Narang, S., Lee, Y. C. & Betenbaugh, M. J. (2004) Comparing N-glycan processing in mammalian cell lines to native and engineered lepidopteran insect cell lines. *Glycoconj J,* **21,** 343-360.

Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. & Berger, I. (2010) New baculovirus expression tools for recombinant protein complex production. *J Struct Biol,* **172,** 45-54.

Tsujita, T., Ishii, A., Tsukada, H., Matsumoto, M., Che, F. S. & Seya, T. (2006) Fish soluble Toll-like receptor (TLR)5 amplifies human TLR5 response via physical binding to flagellin. *Vaccine,* **24,** 2193-2199.

Urlaub, G. & Chasin, L. A. (1980) Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc Natl Acad Sci U S A,* **77,** 4216-4220.

Van Craenenbroeck, K., Vanhoenacker, P. & Haegeman, G. (2000) Episomal vectors for gene expression in mammalian cells. *Eur J Biochem,* **267,** 5665-5678.

Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. & McCawley, P. (1977) The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). *In Vitro,* **13,** 213-217.

Verhoeyen, E., Hauser, H. & Wirth, D. (2001) Evaluation of retroviral vector design in defined chromosomal loci by Flp-mediated cassette replacement. *Hum Gene Ther,* **12,** 933-944.

Verma, R., Boleti, E. & George, A. J. T. (1998) Antibody engineering: Comparison of bacterial, yeast, insect and mammalian expression systems. *J Immunol Methods,* **216,** 165-181.

Volkman, L. E. & Goldsmith, P. A. (1985) Mechanism of neutralization of budded Autographa californica nuclear polyhedrosis virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis. *Virology,* **143,** 185-195.

Volkman, L. E. & Summers, M. D. (1977) Autographa californica nuclear polyhedrosis virus: comparative infectivity of the occluded, alkali-liberated, and nonoccluded forms. *J Invertebr Pathol,* **30,** 102-103.

Volkman, L. E., Summers, M. D. & Hsieh, C. H. (1976) Occluded and nonoccluded nuclear polyhedrosis virus grown in Trichoplusia ni: comparative neutralization comparative infectivity, and in vitro growth studies. *J Virol,* **19,** 820-832.

Wasilko, D. J., Lee, S. E., Stutzman-Engwall, K. J., Reitz, B. A., Emmons, T. L., Mathis, K. J., Bienkowski, M. J., Tomasselli, A. G. & Fischer, H. D. (2009) The titerless infected-cells preservation and scale-up (TIPS) method for large-scale production of NO-sensitive human soluble guanylate cyclase (sGC) from insect cells infected with recombinant baculovirus. *Protein Expr Purif,* **65,** 122-132.

West, A. G. & Fraser, P. (2005) Remote control of gene transcription. *Hum Mol Genet,* **14 Spec No 1,** R101-111.

Wickham, T. J. & Nemerow, G. R. (1993) Optimization of Growth Methods and Recombinant Protein Production in BTI-Tn-5B1-4 Insect Cells Using the Baculovirus Expression System. *Biotechnol Prog,* **9,** 25-30.

Widow, U., Heidler, T., Schmelz, S., Zimmer, A., Kuklik, N., Polle, L., Strube, C., Wengler, U., Wiesand, U., Rochon, M. & Rand, K. (2012) Capability for motivation and exhilaration of depleted doctoral students. Helmholtz-Centre for Infection Research, Braunschweig.

Wilke, S. (2011) Production of mammalian glycoproteins for structural analysis: site-specific recombination systems in CHO cells, **PhD Thesis**, Technische Universität Braunschweig, Helmholtz-Zentrum für Infektionsforschung, Braunschweig.

Wilke, S., Groebe, L., Maffenbeier, V., Jäger, V., Gossen, M., Josewski, J., Duda, A., Polle, L., Owens, R. J., Wirth, D., Heinz, D. W., van den Heuvel, J. & Büssow, K. (2011) Streamlining Homogeneous Glycoprotein Production for Biophysical and Structural Applications by Targeted Cell Line Development. *PLoS ONE,* **6,** e27829.

Williams, G. V., Rohel, D. Z., Kuzio, J. & Faulkner, P. (1989) A cytopathological investigation of Autographa californica nuclear polyhedrosis virus p10 gene function using insertion/deletion mutants. *J Gen Virol,* **70 (Pt 1),** 187-202.

Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U., Schucht, R. & Hauser, H. (2007) Road to precision: recombinase-based targeting technologies for genome engineering. *Curr Opin Biotechnol,* **18,** 411-419.

Wu, S., Hong, F., Gewirth, D., Guo, B., Liu, B. & Li, Z. (2012) The Molecular Chaperone gp96/GRP94 Interacts with Toll-like Receptors and Integrins via Its C-terminal Hydrophobic Domain. *J Biol Chem,* **287,** 6735-6742.

Wurm, F. M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol,* **22,** 1393-1398.

Wurm, F. M., Gwinn, K. A. & Kingston, R. E. (1986) Inducible overproduction of the mouse c-myc protein in mammalian cells. *Proc Natl Acad Sci U S A,* **83,** 5414-5418.

Wurm, F. M. & Hacker, D. (2011) First CHO genome. *Nat Biotechnol,* **29,** 718-720.

Xing, Z., Kenty, B. M., Li, Z. J. & Lee, S. S. (2009) Scale-up analysis for a CHO cell culture process in large-scale bioreactors. *Biotechnol Bioeng,* **103,** 733-746.

Yang, Y., Biedendieck, R., Wang, W., Gamer, M., Malten, M., Jahn, D. & Deckwer, W. D. (2006) High yield recombinant penicillin G amidase production and export into the growth medium using Bacillus megaterium. *Microb Cell Fact,* **5,** 36.

Ye, L., Lin, J., Sun, Y., Bennouna, S., Lo, M., Wu, Q., Bu, Z., Pulendran, B., Compans, R. W. & Yang, C. (2006) Ebola virus-like particles produced in insect cells exhibit dendritic cell stimulating activity and induce neutralizing antibodies. *Virology,* **351,** 260-270.

Yoon, S. I., Kurnasov, O., Natarajan, V., Hong, M., Gudkov, A. V., Osterman, A. L. & Wilson, I. A. (2012) Structural basis of TLR5-flagellin recognition and signaling. *Science,* **335,** 859-864.

Yumerefendi, H., Tarendeau, F., Mas, P. J. & Hart, D. J. (2010) ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *J Struct Biol,* **172,** 66-74.

Zuberbuhler, K., Palumbo, A., Bacci, C., Giovannoni, L., Sommavilla, R., Kaspar, M., Trachsel, E. & Neri, D. (2009) A general method for the selection of high-level scFv and IgG antibody expression by stably transfected mammalian cells. *Protein Eng Des Sel,* **22,** 169-174.

Appendix: Oligonucleotides

Vector construction and mutagenesis

Name		Sequence 5′ → 3′
hr5-ie1-p10_BbsI	for	AGGAGGAAGACTACGGCTGATCATGGAGATAATTAAAATG
hr5-ie1-p10-NsiI	rev	AGGAGATGCATTGTGGATCCGCGCCCGATGG
pFBie10-mut_Bbs-Aat	for	GCTTCAGGAGATCGGACGTCCTGATCTGATCATGG
pFBie10-mut_Bbs-Aat	rev	CCATGATCAGATCAGGACGTCCGATCTCCTGAAGC
Frt-sighis-PGK_BamHI	for	AGGAGGGATCCGGATCGGGAGATCAGC
Frt-sighis-PGK_AvrII	rev	AGGAGCCTAGGCGAGAAGTTCCTATACTTTCTAG
Triex_Cass_SmaI	for	AGGACCCGGGTCGACTATTAATAGTAATCAATTACGGGGTC
Triex_Cass_BamHI	rev	AGGAGGATCCTGATTGTAAATAAAATGTAATTTACAG
pTT_Nco_mut	for	GGTGGGCTAATGTTGCCTTGGGTAGCATATACTACCC
pTT_Nco_mut	rev	GGGTAGTATATGCTACCCAAGGCAACATTAGCCCACC
pTT_Bbs_mut	for	AGGAGAAGACGACCTTTCCTGTTCAAGAATTAATTCTCATGTTTGACAGC
pTT_Xma	rev	GGGTAGCATATGCTTCCCGGGTAGTAG
pTT5-backbone_EcoRV	for	AGGAGATATCAGCGCTGTAGGTGGGCGGGCCAAGATAGG
pTT5-backbone_SapI	rev	AGGAGCTCTTCCGCTATGTGAGCAAAAGGCCAGCAAAAGG
FRT-Cass_BcI	for	AGAGATGATCAGAAGTTCCTATTCCGAAGTTCCTATTCTTC
FRT-Cass_Not	rev	AGAGAGCGGCCGCGAAGTTCCTATACTTTCTAGAGAATAGGAAC

Sequencing primers

Name		Sequence 5′ → 3′
pFlpBtM Seq	for	CGGATCGGGAGATCAGCTTGAAG
pFlpBtM Seq	rev	CAGCCATACCACATTTGTAGAGG
pFastbac Seq	for	TCCGGATTATTCATACCGTCCC
pFastbac Seq	rev	CCTCTACAAATGTGGTATGGCTG
PCR pFast F	for	CCCAGTCACGACGTTGTAAAACG
PCR pFast R	rev	AGCGGATAACAATTTCACACAGG
pFastbac PromSeq	for	CGAACGAGCGCAAGGTTTC
pFS_Seq	for	GGAGATCAGCTTGAAGTTCCTATTCCGAAG
pTT5_Seq	for	GGCCATACACTTGAGTGACAATGAC
pTT5_Seq (pYD5-AS)	rev	TGTCCTTCCGAGTGAGAGACACAA
pTT-mut_seq	rev	AGGGGTTCCGCGCACATTTCC
Seq_Amp	rev	GGGGCGAAAACTCTCAAGG
Frt_Seq_1	for	GAACAACCAAGTGACTCCGGG
Frt_Seq_2	for	CAGGTTCAGGGGGAGGTGT
Frt_Seq_3	rev	AACTGGGTGTAGCGTCGTAAGC
IE1 Promotor	for	TGGATATTGTTTCAGTTGCAAG
IE1 Terminator	rev	CAACAACGGCCCCTCGATA
Seq rtTLR5-c-Term	for	CCTTACAGCGAACCTCATCCA
Seq rtTLR5-c-Term	rev	TCGAACTGCGGGTGGCTC
TLR5_Seq_internal	for	GCCTCGCAGCTAATAGCTTG
eGFP_Prom	rev	CTTGTGGCCGTTTACGTCG

Insert cloning

Name		Sequence 5′ → 3′
tomato_BstBI	for	GGAACTTCGAATTTCCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGG
tomato_BsaBI	rev	GGCTGATTATGATCAGTTATCTACTTGTACAGCTCGTCCATGCCGTAC
BsmBI_mCherry	for	AGGACGTCTCACATGGTGAGCAAGGGCGAGGAGGAT
BsmBI_mCherry	rev	TCCTCGTCTCAGATGCTTGTACAGCTCGTCCATGCCGCC
His_MluI	rev	AAATACGCGTTCAGTGGTGGTG
His_Not	rev	AGGAGCGGCCGCTCAGTGGTGGTGATGATGATG
BsmBI_GFP	for	AGGACGTCTCACATGAGCAAAGGAGAAGAACTTTTCACTGG
BsmBI_GFP	rev	TCCTCGTCTCAGATGTTTGTAGAGCTCATCCATGCCATG
Nco_eGFP	for	CCGGTCGCCACCATGGTGAGC
BsmBI_eGFP_stop	rev	TCCTCGTCTCAGATGCTAGATCCGGTGGATCTGAGTCC
BsmBI_eGFP_His	rev	TCCTCGTCTCAGATGGATCCGGTGGATCTGAGTCCGG
mCherry_Kpn	for	CCCCGGTACCTTATGGTGAGCAAGGGCGAGGAGG
mCherry_Not	rev	CCCCGCGGCCGCCTACTTGTACAGCTCGTCCATGCCG
GFP_Kpn	for	CCCCGGTACCGGCCCATCACAAGTTTGTACAAAAAAGCAGGC
GFP_Not	rev	CCCCGCGGCCGCTCATTTGTAGAGCTCATCCATGCC
GFP_CciI	for	GGTCTATCATGAGCAGCAGCCATCATC
GFP_Not	rev	CCCCGCGGCCGCTCATTTGTAGAGCTCATCCATGCC
mCherry_CciI	for	CGCCATCATGAGGAGCAAGGGCGAGGAG
mCherry_Not	rev	CCCCGCGGCCGCCTACTTGTACAGCTCGTCCATGCCG
hTLR5	for	CCCAGGTCCTCAACACCACTGA
hTLR5	rev	GGTCTTCGTGGCTCACGTCCA
Nco_TLR	for	AGGACCATGGGAGACCACCTGGACC
BsmBI_TLR-wt	rev	TCCTCGTCTCAGATGCTTTAGGGACTTTAAGACTTCCTCTTC
BsmBI_TLR-rtC	rev	TCCTCGTCTCAGATGTATTTTTGTCAAAGATGGTTTTGGGTACTTTTCCG
TLR5rtC_pFlpBtM-I_Nco	for	TACGTCGCCATGGGAGACCACCTGGACCT
TLR5rtC_pFlpBtM-I_BsmBI	for	AGCGTCTCACTCCATTCCTTCCTGCTCCTTTGATGGCCG
TLR5rtC_pFlpBtM-I_BsmBI	rev	TCCTCGTCTCAGATGTATTTTTGTCAAAGATGGTTTTGGGTACTTTTCCG
hTLR5rtC_pFastBac_SalI	for	AGGAGGTCGACATGGGAGACCACCTGGACCTTCTCC
hTLR5rtC_pFastBac_NotI	rev	GCTCTGCGGCCGCTCAATGGTG
ssTLR5_pTT_ApaI	for	AGGAGGGCCCATGGGAGACCACCTGGACCTTC
ssTLR5_pTT_Not	rev	GAGCTCTGCGGCCGCTCAATG
ssTLR5_pTT_Not	rev	CTCTGCGGCCGCTCAATGGTGATG
TLR-wt_QC_BsbBI	rev	TCCTCGTCTCAGATGCTTTAGGGACTTTAAGACTTCCTCTTCATCAC
ssTLR_BamNco	for	AGGAGGATCCAACCATGGGAGATCATCTTCACCTTCTCCTAGGAGTGGTG
TLR-576-Nhe	rev	AGGAGCTAGCTATATCCAAGACACTAAGTGATACAAATACATCAGG
TLR-555-Nhe	rev	AGGAGCTAGCTATGTCCAGGATCTCTAAATTAGCAGG
TLR-533-Nhe	rev	AGGAGCTAGCGAGGCTTAGTCCCCTTAATGCAG
TLR-509-Nhe	rev	AGGAGCTAGCCAAATACAGAACTTGAAGATGAGAAAGTCCC
TLR-480-Nhe	rev	AGGAGCTAGCAAGGAAAAGCTGTTCTAAGCTGGG
TLR-198-Nhe	rev	AGGAGCTAGCCCCTTGTAGGGGCTCGAGCTCATG
ssTLR_BamNco_265_f	for	AGGAGGATCCAACCATGGGTGCCGGGTTTGGCTTCCATAAC
TLR_Nhe_591	rev	AGGAGCTAGCGATAAAAGTGCTAAGTTCACATTCACAAATGAAC
mTLR2H8_pFlpBtM-I_BsaI	for	AGGAGGTCTCACATGCTACGAGCTCTTTGGCTC
mTLR2H8_pFlpBtM-I_BsaI	rev	AGGAGGTCTCTGATGGTGCTGGTGACATTCC
mTLR2H8_pFlpBtM-II_XhoI	rev	AGGACTCGAGCTGGTGACATTCCAAGACG
mTLR2H8 in pTT_XbaI	for	AGGATCTAGACCAATGCTACGAGCTCTTTGGCTC
mTLR2H8 in pTT_NotI	rev	CGAAAGCGGCCGCGTTATCAATGG
scFv-hIgG1Fc_EcoRI	for	AGGAGAATTCCACCATGGGATGGAGC
scFv-hIgG1Fc_BamHI	rev	CTAGAGGATCCTTATTTACCCGGGGAC

Danksagung

Zunächst möchte ich meinem Doktorvater Prof. Dr. Stefan Dübel meinen herzlichen Dank aussprechen. Seine kurzfristige und nicht selbstverständliche Bereitschaft meine Promotion als Mentor zu betreuen weiß ich sehr zu schätzen. Ebenso möchte ich mich bei Prof. Dr. Christiane Ritter und PD Dr. Michael Hust für die Übernahme des Korreferates und des Prüfungsvorsitzes bedanken.

Herzlich bedanken möchte ich mich auch und vor Allem bei Joop van den Heuvel, der mir nicht nur ohne zu zögern die Gelegenheit gegeben hat, in seiner Arbeitsgruppe meinem Wunsch entsprechend in diesem spannenden Thema zu promovieren, sondern mich auch durchgehend exzellent betreut hat. Trotz seines legendären Zeitmangels hatte er immer großes Interesse an meiner Arbeit, war jederzeit für Fragen erreichbar und hatte für jedes noch so kleine Problem einen Lösungsvorschlag. Ebenso hat er es mir auch ermöglicht, an einer ganzen Reihe äußerst interessanter Konferenzen teilzunehmen. Dass er mir zugetraut hat, neben der Entwicklung und Etablierung von mHost-XS auch noch ohne Kooperationspartner gleich mehrere äußerst ambitionierte strukturbiologische Projekte zu bearbeiten, schmeichelt mir. Ich hätte publikationsmäßig wohl ausgesorgt, wenn ich hier jeweils vorne gestanden hätte: (Chen *et al.*, 2010; Yoon *et al.*, 2012).

Andrea Scrima und Bahar Baser gebührt mein Dank an dieser Stelle für ihre Hilfsbereitschaft bei der Durchsicht der Dissertation.

Ganz besonders und von ganzem Herzen möchte ich mich bei „meinen" Studenten bedanken, die die Projekte in Nachwuchsgruppengröße begleitet haben: allen voran bei Carmen Lorenz, für ihre „sehr geschätzte langjährige Mitarbeit in der AG AFF" und ihre unbezahlbare Akribie (ach was: Pedanterie) beim Korrekturlesen von Postern, Papern und der Dissertation. Ebenso bei Gundi Sprick, Mona Wördehoff, Robert Pluschzick, Antonia Kiechle und sogar bei Ewald, dessen Betreuung mir gezeigt hat, was Personalverantwortung auch bedeuten kann. Mangels Kooperationspartner war ihre Arbeit ein wichtiger Beitrag zum Gelingen der Projekte.

Meinen Dank möchte ich auch allen anderen Kollegen der Rekombinanten Proteinexpression aussprechen, welche die Projekte tatkräftig durch praktische Hilfe oder nützliche Tipps unterstützt haben: den Mitdoktoranden Bahar, Jens, Sonja und Zhaopeng sowie den TAs Nadine, Daniela, Sarah, Steffi, Claudia und Anke. Konrad Büssow und Volker Jäger, die jederzeit ein offenes Ohr für meine Fragen hatten, danke ich zudem auch für ihre hilfreiche Diskussionsbereitschaft.

Bei den vielen anderen Kollegen und Ehemaligen der Strukturbiologie (Widow *et al.*, 2012), die ich *an dieser Stelle* nicht alle einzeln aufzählen kann, danke ich dafür, dass sie die Arbeit hier über die Jahre so vergnüglich gemacht haben. Ein Betriebsausflug ohne Ute wäre möglich, aber sinnlos.

Für das große Maß an Motivation, ohne das eine langjährige naturwissenschaftliche Doktorarbeit nicht möglich wäre, haben die vielen Freunde in der gemeinsamen Zeit außerhalb des Labors gesorgt, bei denen ich mich dafür auch herzlich bedanke. Und nicht zuletzt natürlich bei Franziska, der besten TA der Welt, für überhaupt Alles.

Lebenslauf

Persönliche Angaben

Steffen Meyer
geb. am 24.3.1982 in Bielefeld
ledig
deutsch

Schulische Ausbildung und Wehrdienst

8/1988 – 7/1992	Marienschule Verl
8/1992 – 6/2001	Gymnasium Verl
11/2001 –7/ 2002	Wehrdienst

Beruflicher und wissenschaftlicher Werdegang

10/2002 – 2/2008	Studium Biotechnologie (Diplom) an der Technischen Universität Carolo-Wilhelmina zu Braunschweig
10/2005 – 9/2006	Studienarbeit am Institut für Mikrobiologie der Technischen Universität Carolo-Wilhelmina zu Braunschweig: „Untersuchung von Promotersystemen in *Bacillus megaterium*"
4/2007 – 2/2008	Diplomarbeit am Helmholtz-Zentrum für Infektionsforschung Arbeitsgruppe Molekulare Wirt-Pathogen Interaktionen: „Strukturelle, biophysikalische und funktionelle Charakterisierung der Wirtsspezifität von *Listeria monocytogenes*"
4/2008 – 9/2008	Trainee Abteilung Process Development, Glycotope GmbH
seit 10/2008	Promotion am Helmholtz-Zentrum für Infektionsforschung Arbeitsgruppe Rekombinante Proteinexpression

Publikationen

Biedendieck, R., Gamer, M., Jaensch, L., Meyer, S., Rohde, M., Deckwer, W. D. & Jahn, D. (2007) A sucrose-inducible promoter system for the intra- and extracellular protein production in *Bacillus megaterium*. J Biotechnol, 132, 426-430.

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. Büssow, K., Jäger, V. & van den Heuvel, J. New Versatile Donor Vector for Protein Production in Eukaryotic Systems, (in submission)

Meyer, S., Baser, B., Sprick, G. Lorenz, C. & van den Heuvel, J., mHost-XS - An integrated versatile expression platform for protein production in eukaryotic hosts (in preparation)

Tagungsbeiträge und Seminarvorträge

Meyer, S., Pluschzick, R., Fischer, A., Hagendorf, P., Stradal, T., van den Heuvel, J., Expression of the multiprotein complex WAVE in insect cells using the MultiBac Expression System, 9th conference on Protein Expression in Animal Cells (9th PEACe), Jackson (Wy) (2009), Poster

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. New Versatile Donor Vectors for Protein Production in Eukaryotic Systems, ISBioTech 1st Annual Meeting, Norfolk, VA (2011), Vortrag

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. A New Versatile Donor Vector for Protein Production in Eukaryotic Systems, SGC Workshop on Human Protein Production, University of Oxford (2011), Vortrag

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. A New Versatile Donor Vector for Protein Production in Eukaryotic Systems, PEGS Europe - Protein & Antibody Engineering Summit, Hannover (2011), Poster & Vortrag

Meyer, S., Lorenz C., Sprick, G., Baser, B. & van den Heuvel, J. A New Versatile Donor-/Expression Vector for Protein Production in Eukaryotic Systems, Animal Cell Technology Industrial Platform (ACTIP) Biannual Meeting, Lyon (2011), Vortrag

Meyer, S., Lorenz C., Baser, B., Wördehoff, M. & van den Heuvel, J. A New Versatile Donor Vector for Protein Production in Eukaryotic Systems, PEGS Europe - Protein & Antibody Engineering Summit, Hannover (2011), Vortrag

Meyer, S., Lorenz C., Sprick, G., Baser, B. & van den Heuvel, J. mHost-XS - A versatile System for Protein production in eukaryotic systems, Seminar on special topics of Molecular and Technical Biochemistry, Technische Universität Braunschweig (2011), Seminarvortrag